



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 142280

To: Barba Koroma
Location: REM/2C18
Art Unit: 1638
Wednesday, January 26, 2005

Case Serial Number: 10/808979

From: Beverly Shears
Location: Remsen Bldg.
RM 1A54
Phone: 571-272-2528

beverly.shears@uspto.gov

Search Notes



This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 23:06:56 ; Search time 28265 Seconds
(without alignments)
11709.910 Million cell updates/sec

Title: US-10-808-979-18
Perfect score: 6999
Sequence: 1 gcgatggaatggaatgt.....gacgttttagacatgcaata 6999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2732.8	39.0	6815	AB076662	AB076662 Pantoea a
2	2626.2	37.5	6918	1 ERWCRT	D90087 Erwinia ure
3	2624.6	37.5	6918	6 I12910	I12910 Sequence 13
4	2620	37.4	6586	1 AY166713	AY166713 Pantoea s
5	2468.8	35.3	6965	1 ERWCRTA	M90698 Pantoea agg
6	2135.2	30.5	12753	1 ERWCRTS	M87280 Pantoea agg
7	1900.6	27.2	4624	6 CQ93208	CQ93208 Sequence
8	1374.8	19.6	349107	1 BX571873	BX571873 Photorhab
9	1189	17.0	2415	1 ERWCRTIB	M38423 Pantoea agg
10	1179.8	16.9	3801	1 AF408848	AF408848 Xanthobac
11	1105.8	15.8	6635	1 AF218415	AF218415 Bradyrhiz
12	1041.2	14.9	6335	1 PMCRTWZYI	Y15112 Paracoccus
13	1037.2	14.8	5373	1 ATUCRTWA	D58420 Paracoccus
C 14	1035.4	14.8	5188	1 FAU62808	U62808 Flavobacter
C 15	1035.4	14.8	8625	6 AR169831	AR169831 Sequence
C 16	1035.4	14.8	8625	6 AR452005	AR452005 Sequence
C 17	1035.4	14.8	11233	6 AR169852	AR169852 Sequence
C 18	1035.4	14.8	11233	6 AR452026	AR452026 Sequence
19	827.4	11.8	1493	6 BD241820	BD241820 Method of

20	827.4	11.8	1493	6	AX014697	AX014697 Sequence
21	826.2	11.8	1479	6	CQ793426	CQ793426 Sequence
22	826.2	11.8	1479	6	CQ793595	CQ793595 Sequence
23	826.2	11.8	1479	6	E22381	E22381 beta-Carote
24	826.2	11.8	1479	6	I12907	I12907 Sequence 10
25	808.6	11.6	1479	6	AX394988	AX394988 Sequence
26	713.4	10.2	1518	6	I72653	I72653 Sequence 7
27	708.4	10.1	1522	6	I72654	I72654 Sequence 9
28	619.4	8.8	1482	6	A58565	A58565 Sequence 6
29	619.4	8.8	1482	6	A84702	A84702 Sequence 12
30	619.4	8.8	1482	6	AR103081	AR103081 Sequence
31	619.4	8.8	1482	6	AR139697	AR139697 Sequence
32	619.4	8.8	1482	6	AR391876	AR391876 Sequence
33	604.4	8.6	349723	1	BX42850	BX42850 Bdellovib
34	515.6	7.4	1198	6	I24838	I24838 Sequence 5
35	515.6	7.4	1198	6	I72652	I72652 Sequence 5
36	479	6.8	30040	1	AF182374	AF182374 Bradyrhiz
37	467.4	6.7	349737	1	BX572597	BX572597 Rhodosphe
38	461.4	6.6	1581	1	D83514	D83514 Erythrobact
39	461.4	6.6	1581	6	E11103	E11103 Gene coding
40	449.6	6.4	1083	6	I40224	I40224 Sequence 1
41	449.6	6.4	5397	12	AY605097	AY605097 Synthetic
42	449.2	6.4	1296	6	AX394984	AX394984 Sequence
43	438	6.3	1235	6	I72656	I72656 Sequence 13
44	430	6.1	1232	6	AR222522	AR222522 Sequence
45	430	6.1	1232	6	AR432935	AR432935 Sequence

ALIGNMENTS

RESULT 1
AB076662
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

AB076662
Pantoea agglomerans pv. milletiae crtX, crtY, crtB, crtZ genes, complete cds.
AB076662.1 GI:18143444
Pantoea agglomerans pv. milletiae
Pantoea agglomerans pv. milletiae
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Pantoea.

Kamiunten,H. and Hirata,R.
Isolation and characterization of carotenoid biosynthesis genes from Pantoea agglomerans pv. milletiae Wist 801
Unpublished
2 (bases 1 to 6815)
Kamiunten,H. and Hirata,R.
Direct Submission
Submitted (20-DEC-2001) Hiroshi Kamiunten, Miyazaki University, Faculty of Agriculture; Gakuenkibanadai Nishi 1-1, Miyazaki, Miyazaki 889-2155, Japan (E-mail: a01108@cc.miyazaki-u.ac.jp, Tel:81-985-58-7170, Fax:81-985-58-7170)
Location/Qualifiers
1. .6815
/organism="Pantoea agglomerans pv. milletiae"
/mol_type="genomic DNA"
/db_xref="taxon:182454"
/note="pathovar: milletiae"
160. .1071
/gene="crtE"
160. .1071
/gene="crtB"
/codon_start=1
/transl_table=11
/protein_id="BAB79600.1"
/db_xref="GI:18143445"
/translation="MMTVCAEOHVFHSDAASLLNDIEQRDLQLPVSERDLVGAAMRDGALAPGKRIRPLLILLARLDLCGNATPAGLLDACAVEMVHAASLLIDDPICMDDQALRGRTTHCQYGEHVAILANVALLSKAFGVAAAEGLTATARADVAELSHAVGMQQLVQGFQKDLSEGDKPRSDAILMTNHYKTKTLFCASMQMASIAVAEASGEAREQLHR

FSNLQAFQLDDLDLTDGMADTKDAHQDDGKSTLVNLLGPQAVETRLRLHLCASEH
 LLSACODGVAATHFVQAWFEKKLAAS"

gene

1081..2376

CDS

/gene="ctx"

1081..2376

/gene="ctx"

/codon_start=1

/transl_table=1

/protein_id="BAB79601.1"

/db_xref="GI:18143446"

/translation="MSHFAVIAPFFYSHVQALQHLISQALIARGHQITFIQOQTVSALL

TSRIGFFPLGLASHPAGSLAHTLOLAHPLGSMKLNEMARSTDMCLRELPPVLS

TLAIDGVIYDQAGALAAEALNLPYVSACALPLNREADPFLAVMPDYARTDQAR

ERYTRSEKIYDLMRRHRDVIARNHAMGLAPREKLHCFSPVAOINOLIPELDPRQ

TLPDHPHSVGLRTPTPAASQSPRYFPHGDRPRIPASLGTIQGHRYGLFKTIIVRACQ

EIDAOLLAHCGRLSPFOAEKLAQASHVQVDFADQAAALAAQADLVIHGMNMTVLDG

INHLPTLLIPIAPOPQVAARVWHGIGRRASRFTTSHMARQJQTLILADESYVQQM

KILRSALRQAGGTTLAADIIVEQAMLTROPVLTFRHYAAV"

2363..3523

/gene="ctx"

2363..3523

/gene="ctx"

/codon_start=1

/transl_table=1

/protein_id="BAB79602.1"

/db_xref="GI:18143447"

/translation="MPRYDLILVAGLANGLIALRLROQPSRLILLIDAEREPCANH

TWSFHAEDTETQHRWIAPLVHVHWPGEVREPORSRNSNGVFCVTAERPAQVIRDR

FAPDILLINTRVIAASRAVTLDDGRVLESDAVIDGRGQPDGALRMGFSFVQGQWL

SPHGUTAPIIMDATVDQAGRYFVSUPFSADTILLIEDHYDNATLEGDRARQNIH

DYAQQWRLLDRLEEGALPITLIGDVAFWQNRDLPCSLGRAGLPHFTTGYSLP

AVALLADRAQMTFTSETLHATIQOFASQAOQORFFRMLNRLMLFLAGPADORWQVQ

RYFGLPEGLIARFYAGKLTLPDLRLILSKPPVPVLAALQALIMTPHRQAMQ"

3520..4998

/gene="ctx1"

3520..4998

/gene="ctx1"

/codon_start=1

/transl_table=1

/protein_id="BAB79603.1"

/db_xref="GI:18143448"

/translation="WNRITVIGAGFGLALAIRLOASGVPTLLEBQDKPGGRATVYQ

DQGFIDPAGTIVTPSAIEELFLTAGKKLSYDELMPVKPYPYRLCSWGVKFSYDND

QPALBAQAFNPVDEPREFLYASRAVFAEYGLKGLTVPFLSRDMLRAAPQLAKL

QAWRSYKVASIYIEHLRQAFSHFLLVGNPATSIYTLIHAREWGVMPVRG

GTGALVQGVKLFIDELGVELVNASVARLEQENRITAVHLKDGVRVTPRAVASNADY

VHTYRELLISOHPASOAGRSLOKEMSNLSFVYFGLAHNHNOLAHHTVCGPRYREL

IDEIIFKGLADEFSLYLHAPCTVDPSPSLAPECCGSYYVLAPVPHLGTADIDWAVEGR

LDRIFDIEQHYMPGLRSQLVTHRIFTPTDFRDELNAYQGSFAFSVEPILTQSAWFRP

HNRDKNNLNLVVGAGTHFGAGIPGIVGSAKATAGLMLDLA"

5034..5924

/gene="ctxB"

5034..5924

/gene="ctxB"

/codon_start=1

/transl_table=1

/protein_id="BAB79604.1"

/db_xref="GI:18143449"

/translation="NEVGSKSPATASLFGAKTRSLVLMVAMCHCDVDIDQVLF

SNDTPSLOSARQALQLEMKYTRQAGSQMHEPAPAAQEVNMAHDILPAYAFDHLG

FAMDVHETRYQLDILRYCHYVAGVGLMMAQIMGVDNATLDORACDLGLAQLTNI

ARDIVEDADAGRCYLPATWLAEGLTRENLDAPQNRQALSRRVRLVETABPYRSAS

AGLGLPLRSAWATYATQOYVRKIGMKVQVQAASQAWDQORQSTSTPEKLALLVAASQA

VTSRVARHAPRSADLWQRPV"

complement(5863..6390)

ORIGIN

Query Match 39.0%; Score 2732.8; DB 1; Length 6815;
 Best Local Similarity 65.8%; Pred. No. 0;
 Matches 4085; Conservative 0; Mismatches 2092; Indels 27; Gaps 7;

QY 376 CTGCGAGCAAAATTTTACAGCGGCATCTTGAACATTTTACTGCCTGCCGGACAGCAAAACGCGAT 435
 Db 217 CTGTTGAAACACATTTGAGCAACCGCTTGATCAGCTTTTACCGTTTGAAGCGAACGCGAC 276
 QY 436 CCGCGTGGTGGCGCGCATCGTCCCGAACGCTGGCGGAGGCAAAACGTTTTCGTCTCTTTA 495
 Db 277 TTATGGGGCGCTGCCATCGCGACCGTGGCTGGCACCAAGGAAACGCTATCGCTCCACATG 336
 QY 496 TTACTGTCTGTGGCAGCGCGCATATGGTGTGCGAGCTGACGCAAAATGGCGTTCTCGAT 555
 Db 337 CTGTTGTTGCTGGCAGCGCGCATCTGGCTGCAACGCCACGCCCTGCCGCTGCTCGAT 396
 QY 556 CTCGCTGTGCAGTGGAAATGGTGACCGCGGATCGCTGATCTTGGATGACATTCCTCG 615
 Db 397 CTCGCTGTGGCGTAGAGATGGTGCATCGCGCATCACTGATTTCTGGATGACATGCGCTGC 456
 QY 616 ATGGATAACGGCAGATCGCTCGTGGTGGTCCCTACCGTGCATCGCGAATTTGGTGAAGAAC 675
 Db 457 ATGNATGATGGCAACTCGTGGCGGACGCTCGACCATTCATTTGCCATGATGGTGAACAT 516
 QY 676 GTGGCGATTTCTGGCGCCATCGCGTGTCTAGCCGCGCATTTTGAAGTGAATTCGCAATGCA 735
 Db 517 GTCGCGATTTCTGGCGCGGTGGCCCTGCTGAGTAGGACATTTCCGGCTGCTCGCTCGCGCA 576
 QY 736 CCCGTTTGCCTGCCATACATAAATCTGAAGGATTTGCTGAACTCTCCGCTGCGCTCGCGC 795
 Db 577 GAAGGCTTTAACCGGCAACCGCCAGAGCCGACGCTGTGGCAGAAATTTATCCACGCGACGCGC 636
 QY 796 CTGCAAGGCTTTAGTGAAGGCAATTTCCAGGATCTGCACGACGCGCACGACGCGCGCAGC 855
 Db 637 ATGCAGGGGCTGTGTGAGGGGCAATTTAAGATCTCTCCGAGGGTGACAAGCCAGCGCAGC 696
 QY 856 CCGAAGCGATTCGCCATGACCAAGAACTGAAAAACCGAGCTGTGTTCGGCGCCACGCTG 915
 Db 697 GCTGACGCCATTTCTGATGACCAATCACTATATAAACACGACGCTGTTCGGCTCCCATG 756
 QY 916 CAATGGCGCGATTTGGCGCTGACGCTTCAACGCGAGGTGGCGCAAGACTTAGCTTCTTC 975
 Db 757 CAGATGGCTCTATCTGTGGCTGAAGCCCTCAGGTGAAGCCCGCGCAACAGCTGCAACCGTTT 816
 QY 976 GCCCAGGATTTGGGCGCGCGCTTTCACTGCTCGACGACCTCGCGCGCGGTTGCAAAACAC 1035
 Db 817 TCGCTTAATCTTGTGTCAGGCTTTCCAGCTACTTGGACGATCTCATGACGGATGCGCGAC 876
 QY 1036 ACCGGTAAAGATGTGACACGAGGATGAGGCAAAATCCACGCTGGTACAGATGCTCGGTGCT 1095
 Db 877 ACCGGTAAAGATGCCATCAGATGAGGGAAATCAACGCTGGTGAATCTGCTGGGACCA 936
 QY 1096 GACGGCGGGAACGCTCGCTGGCGGATCACTCGCGCAGCGCAGATGACACCTTGCCTGC 1155
 Db 937 CAGCGGTGTAAGACGCGACTCGCGGATCATCTGCGCTGCGCGCAGCGACATCTGTTATCG 996
 QY 1156 GCCTGCCATCGCGCATCGCCACTCGCAATATATGACGCGCTGTTTAAATCAACAGCTA 1215
 Db 997 GCCTGCCAGGACGCTTATGCCACACACCAATTTGTTGAGGCTGTTTGAAGAAAATC 1056
 QY 1216 GCGATATTCAATGAGCGCGCTCAGCGGTGGGCGCATTTTGGGCTGATCGCGCGCGCGC 1275
 Db 1057 GCTCGCGTCAAGGATGCTGCA----TGAGCCCACTTTTGGCGGTCAATTGACACCGCGCT 1111
 QY 1276 TCTACAGCACTTTTCAAGGCTGAGGCGTTAGCACAAAACGCTGCTCGCGCGCGCGCAATC 1335
 Db 1112 TCTACAGCACTGTCAGGCGCTTCAGCACTTACCGAGGCTTAATCGACGCGGACATC 1171

1336 GCATCACATTCCAGCAGCCGATGCCCGCACTTTGCTTAGCGACGAACGCATCGATT 1395
1172 AGATTACTTTTCATCCAGCAGCGAGCTTAGCGCGCTACTACCGATAGCCGTATCGGCT 1231
1396 TTGTTGCCGTCGCGCAACAGACGCACTCCCTGCCGCTTCCGCTGCGCGCGGCTGTCATCGGC 1455
1232 TTTTCCCGCTGGGTTTAGCCTCGCATCCGCGCGGCGAGTCTGGCACACACCCTGCAACTGG 1291
1456 TGGCCTCGCGCGCGCGCTGTCGCTGTTTTCGCGTGTGATCGACGATCTCGGCTCTGACCG 1515
1292 CGCGCATCTCTCGCGCCCTCAATGCTCAAACTGATCAATGAGATGGCCGCGCAGCACCG 1351
1516 ATATGCTGTCGCGCAACTGCTCGGTACTGAAGCACTTGAACATTCGATGCGGTGATCG 1575
1352 ATATGCTGTCGTTGAACCTGCCGCTGCTGAGCACCTTGGCGATTGATGGCGTATG 1411
1576 CCGACGAATGAAGCGCGCGCGGATTCGCTGAAAGCGTGCATCTGCCGCTTGTGTT 1635
1412 TCGATCAGATGAGCCTCGCGGTGCATCGCGCGAGAGCGCTGAACCTGCCTTATGTTT 1471
1636 CGGTGCTGCGCCTTTCGCGTCAATCGTGAAGCGGATTCGCTTGGCGTGTGATGCCCT 1695
1472 CGGTTGCTGTCGCTGCCACTTAACCGTGAAGCGGATTTCCGCTGCGCGTGTGATGCCCT 1531
1696 TCGGTTTTCACAGGATGACAAAGCGCTGAAGCGTTTTCAGGCCAGCAGCATATCTATG 1755
1532 TTGACTATGCCAGGACCGATCAGGCACGGAACGCTATCGCACCACTGAAATACTATG 1591
1756 ATCGATCATGCTGCTCAGCGCAGCTGATCTCAACACGCGCGGCGTTTAATTTGA 1815
1592 ACTGGCTGATGCGATGACGATCGGTGATCGCCGCAACGCTCATGGCGATGGGATTAG 1651
1816 CGGAGCGCGCGGATTAATCAGTGTCTGCGCGCTGCGCAAAATCAGCCAGATGGTGC 1875
1652 CACCACGGAAAACTGCATCACTGTTTTCACCGGTGGCGCAGATCAACAGCTGATTTC 1711
1876 CGGCTTTGATTTTCCAGCTCAGCACTGCGCGCTGCTATCACGCGTGGCGCACTCC 1935
1712 CGGAATGATTTTCCCGCCACAGACACTGCCCGATCAITTTCCATTCGCTTGGCGCGTGC 1771
1936 GCGCC-----CGGTTTCTCTGCGCGCTCCATGCGCGCTTGGCGCGCTGCGTCAAG 1989
1772 GTACCACTGAACCGGCACCGCGCATCTCAGCTCGCTACTTCCGCACTGCGCAGAC 1831
1990 CGGTGCTTTATGCTCGCTGGGTACGCTGCAAGGCGATCGCTTCCGCTTCTTTCGATC 2049
1832 CGCGCATTTTGTCTCACTCGGCAGCTGCGAGGACATCGCTAGCGCTGTTTAAACCA 1891
2050 TGGCGCAGCGTGCAGCTGCGGCTATCGCTGTGATGCCCATTTGTGGGGATTAA 2109
1892 TCGTCCGCGCTGCCAGGAGATCGATGCTCAGTTATTGCTGGCGCATTCGCTGCGTGTG 1951
2110 ACGCGAAGCAGCGCATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGGATTTGCTCG 2169
1952 CACCTTTCCAGCGGAGAACTGGCGCAGCGCCAGCATGTTTCAAGTGTGCTGATTTGCGC 2011
2170 ATCAGCGCAGCCTTACAGCAGCGCAGCTGTTTATCACTCATGCGCGGTTTAAACAGCG 2229
2012 ACCAGCGCTGCACCTGGCGCAGCGCGATCTTGTCTATTAGCAGCGCGCATGAATACGG 2071
2230 CGGTGAAGCACTGGAATCGGGTACCGCGATGCTGCGCTGCGGATTTGCTGATCAGC 2289
2072 TGTGATGGGATTAACCACTGACCGCTGCTGACGATCCGCTGCGCTTGTGATCAGC 2131
2290 CGGCGTGGCGCGCATTTGATGGCATGAGCTTGGTCGCGCGCATCAGCTTTTAGCC 2349
2132 CGGCTGGCTGCCAGAGTGTCTGGCAGCGTATCGAGTTCGCGCTCAGCTTTTACCA 2191
2350 GTGTTTCATCACTGAGCAGCATCTGCAACAGCTGCTGACCGCAGATCGTTAGCGCTAC 2409
2192 CCAGCCACTCCATGGCAGCTGACCTCCAGACCTTACTGGCTGATGAAGTACGTTTACG 2251
2410 GGATGTACGGATTACGGCGCAGCTGACGCGCGCGGCTTGCAGCGTGGCGCGCAGCA 2469

2252 AGATGAACCGCTGCTAGCGCATTTGTCAGCGCGGCGGAACCACTTGGCGCGGACA 2311
2470 TCGTCAGCAGCGCGCTGTGCCAGCAGCAAGTCTGCTGCGGAGCGGACCTGATCGCAC 2529
2312 TCGTCGAACAGCGGATGCTGACCCGCCAGCGGTCTCAC-CAGGAGACACTATGC--C 2367
2530 GCATACAGATGATGATTTTGGTCTGGATCTGGGATGCGCTGATGCTGATGCTGCTGCT 2589
2368 CGGCTATGATCTGATTTCTGGTAGGTGCGGACTTGGCAACCGGCTGATCGCTTGGCT 2427
2590 GGGTCAATTGACGCCACAACTGAAATGCTGTGCTGAGAGCGATGCGCATCGGCGAG 2649
2428 GAGACAGCAGCGGCCCTCCCTGCGCATTTCTGCTGATTTGACGCGGAACGTAACCGGCTG 2487
2650 CAATCATACCTGGTCTGTTTCATCAGCGGATCTCAGCGCGCAACAACTTCTGCTGGCTGCA 2709
2488 CAATCATACCTGGTCTGTTTCATGCGGAAGATCTCACCGAAACGAGCATCGCTGGATCG 2547
2710 ACCGCTGATACCGTGGCTGCTGAGTATCAGGTGCGTTCCTGCGCTGCGCGCA 2769
2548 TCGCTGCTAGTACATCACTGGCTGCTGATGAGGTCCGCTTTCCCAACGAGTCCGAG 2607
2770 TCTGACCGGGATTAATGTTCCATCGCATCAGCGGATTTTGGCCGCGCATCTTTACCGCG 2829
2608 TATGAACAGTGGCTATTTTGGTGACTCGCGAGCGCTTCGCGCAGGTCTATCCGCGAC 2667
2830 GATGGTGAAGTCTGTGGACAAACACAGCCGTAACAAGGTAAACCCACGAGGTGAC 2889
2668 GTTTCGCGCGATCTGCTGCTGAATACCGGGTTCGAGCATCGCTTTCACGCGCGTCA 2727
2890 GCTGCGGATGCGCGTGAACCTTGTGCGCAAGTGTGATGATGCTGCGCGCTGCGAC 2949
2728 GCTGACGATGCGCGGCTGCTGGAGATGACAGTGTGATGATGCTGCGCGTACAGCC 2787
2950 GAGCGCACTCTGACAGCTGGGTTATCAGGTGTTTCTTGAACAAGTGTGAGTGGCGCA 3009
2788 CGATGCGCGCTGCGCATGGGCTTCAGTCTGTTTTCGAGCAGGATGCGAGTCTCAGCG 2847
3010 GCCGCAAGCGCTGACAGCAGCGATCTGATGATGATGATGATGATGATGATGATGATG 3069
2848 GCCACATGSCCTGACCGCGCGCATCATCATGATGATGATGATGATGATGATGATGATG 2907
3070 TCGTTTGTCTACAGCTGCGCTCAGCGCGATCGGCTATTTGATTTGAAGATACCCATTA 3129
2908 TCGCTTTGTCTACAGCTGCGCTTTTCTGACAGCAACCTGCTGATTTGAAGATACCCACT 2967
3130 CGTTAACAGCGCGCTGCGCGAGAACACCGCTGCTGACGACATCGCGAGCTATGCCAA 3189
2968 CATGATTAACGCGCATGGAAGGCGATCGCGCGCTCAGATATCCAGCATATGCGCG 3027
3190 TCAGCAAGCTGACGCTGAGTACGCTGCTGAGAGCAGCGCATATTTACCGATTAC 3249
3028 GCAGCAGGATGCGATTAGACCGGCTGCTACGGAAGAGCAAGCGCTTACCCATTAC 3087
3250 CTTGAGCGCAACATCGATCGATTCTTGGCAACAGCAGCGCGCGCAAGCTGAGCGCGCT 3309
3088 GCTGACCGCGCAGCTGGCGCGCTTCTGG---CAGAACCGCGATTTGCGCTGCGAGCGT 3144
3310 GCGCGCGCGCTGTTTCATGCGCACACCGGTTTCTCTTGGCTGCGCGCTGCGCGTACG 3369
3145 GCGTGGCGGTTTATTCATCCCACTACCGGCTATTTCTGTCACACTGCGCGGTTGCTTGG 3204
3370 GGAGTTGCTAGCAGCGCTTTCGCCCAACCGATCCCTCAGCTCAGCGCAACATATCGAAC 3429
3205 TGACCGCTGGCGAGATGACAGCTTTTACCTCTGAGACCTTGCAGCCCACTTACGCA 3264
3430 CTTTGGCGCTGACAGTGGCGCAACAGCGATTTTTCGCTTCTGCTAAACCGCATGCTGTT 3489
3265 ATTGCGAGTCAAGCGCTGGCAGCAACAGCGCTTCTTCCGATGCTGAAACGATGCTGTT 3324
3490 TTTGGCGGTAGCGCAGCGCTGCGCTGATGCAACGTTTTCGCTGCGCTCGATGCTGCG 3549

Db 3325 TCTGGCGGCTCCGGCTACACAGCGCTGGCAGGTTATGCAACGTTTTTACGGCCTTCCCGA 3384
QY 3550 CGGTTAAATTTAGCCGCTTTTACGCGCGCAACTGCGCTGCGGATAAACCGGATTTCT 3609
Db 3385 AGTCTGATCGCCGCTTTTATGCGGAAACTCACTTTGCTGATCGGCTGCGCATCTT 3444
QY 3610 GTGCGCAAGCCCGCGTGCCTATCGGTGAAGCGCTGCGCGCTCTTTGAATTTCTGCGA 3669
Db 3445 AAGCGCAAGCCCGCTCCCGTTCTGGCGGTTTACAGGCTATTATGACTCTCCACCG 3504
QY 3670 ACCAGGAGAAAAATGAAACGCACTTATGTAATTTGGCGCAGGCTTTGGCGCTGGCG 3729
Db 3505 TCAACAGGCGATGCAATGAAATAGAACTACAGTAATTTGGCGCAGGCTTTGGTGTCTGGCT 3564
QY 3730 CTGGCAATTCGCTCGCAAGCGGCGCATACCAACCACTTACTCGAGCAGCGGACAAA 3789
Db 3565 CTGGCAATTCGCTTCAGGCTCAGCGTTCCACCCGACTGCTGGAGCAGCGTGACAAG 3624
QY 3790 CCGGCGGACGCGCTATGTTTGAAGACAGTGGCTTTACCTTCGATCCCGACCCACG 3849
Db 3625 CTTGGCGCGCGGCTTATGCTATCAGGATCAGGGCTTCACGTTTGATCGCGGCCACG 3684
QY 3850 GTGATCAGCATCCAGCGCCATCGAAGATTTGTTTCACTGCGCAGGAAAAATCGCTCAGC 3909
Db 3685 GTAATCACCAGTCCCGAGCGCATTTGAAGAGCTGTTTCACTTCGCGGTTAAAAAGCTCTCT 3744
QY 3910 GATTACGTCAGCTGATCGCGTAACGCTTCTATCGGCTGTGCTGGGAAGATGGCAA 3969
Db 3745 GACTATGTCAGCTGATCGCGTGAAGCGTTTATCGGCTCTGCTGGAGTCCCGCAAG 3804
QY 3970 CAGCTTGATAGCAATTAATCAGCGCTGCTGGAGCAGACATCGCCAGCTTCAATCCG 4029
Db 3805 GTGTTCAAGTTATGACAAAGATCAGCGCGCTGGAAGCGAGATGCGCGGTTAATCCG 3864
QY 4030 CAAGATGTAAGAGCTACGTCATTTCTTGCTTATTCAGTGAAGTATTAGAGAGGT 4089
Db 3865 CTGAGCTTGAAGGATATCGTCGCTTTTGGGCTATTTCCGAGCGGTGTTTCTGAAGCG 3924
QY 4090 TATCTGAAGACTCGGACGCGTCCGTTCTGCAAGTGGCTGACATGCTGCGCTGCGCGC 4149
Db 3925 TATCTGAAGCTTGGCACCGTCCGTTCTGTATTCGCGCAATGCTGCGCGCGCGCT 3984
QY 4150 CAGTTGGAGCTGTGCAAGCATGCGCAGCGTCTACAGCATGCTGGCGAAATTTATTCAG 4209
Db 3985 CAGCTGGCAAAATTTTCAAGCGTGGCGCAGCGTTTACAGCAAAAGTGGCGAGCTACATGAA 4044
QY 4210 GACGATCATCTGCTCAGCGCTTTCTTCCACTCATTTGCTGCTGGCGGTAATCTTTT 4269
Db 4045 GATGAGCATCTGCTCAGGCTTCTTTCCACTCACTGCTGGTGGGGAATCCGTTT 4104
QY 4270 GCAAGCTCATCATCTATACCTTAATTCATGCGCTGGAGCGTGAATGGGCGTGTGTTT 4329
Db 4105 GCCACTTCTCAATCTATACCTGATTCATGCTGGAACGTTGATGGGCGTCTGGTTC 4164
QY 4330 CCGCGGCGGACACCGCGCGCTGTGCAAGCATGCGCGCATGCTTTCGAGACTTGGGC 4389
Db 4165 CCGCGCGGTGGACCGCGCGCTGTGTCAGGCGCATGGTGAATCTGTTGAAGATCTGGGC 4224
QY 4390 GCGGAGCTGTTACTGAATGCCGAAGTGAAGCAGCTGGAACCCAGCGCAATCGCATTTAGC 4449
Db 4225 GCGGAAGTGGAGCTCAATGCCAGGTTGCGCGGCTGGAGACCCAGGAATAAGAAATTACC 4284
QY 4450 GCGCTTCAGTTAGGGCGGACGACGCTTCGATGCGCGCGCTGTGGCTTCCAAATCCGAC 4509
Db 4285 GCGGTGCACTGAAAGATGGCGGCTTCTCCGACCCCGCGGTTGCTCCAAACGAGAT 4344
QY 4510 GTGTGTCATACCTACGACAACTGCTTCGCCACCATCGCTGGCAATGAACGTCGACA 4569
Db 4345 GTGGTTACACCTACCGGAACTGCTGAGCCAGCATCCCGCTTCGACGCGCGGAGCA 4404
QY 4570 TCGCTGAAGCGTAAGCGATGAGCAACTCGCTGTTTGTACTTATTTTGGCTGTAATCAG 4629
Db 4405 TCACCTGCAAAACAAACGATGAGCAACTCACTGTTTGTGATCTATTTTGGCTGATCAT 4464

QY 4630 CCGCATGAACAGCTCGCGCACACACCGCTCTCTTTTGGCCCCGCTTATCGTGAAGTTGATC 4689
Db 4465 CATCAAAATCAGCTGGCGCACACACCGCTCTCTTTTGGTTCGCGCTATCGTGAAGTTGAT 4524
QY 4690 GATGAGATTTTCAACAGCAGCGCAGCTGCGCAGACGATTTTCACTTTTCACTGACGCGCC 4749
Db 4525 GATGAGATCTTTAAACAGATGGCTTGGCAGAGGACTTCTCGCTCTATCTGATGCGGCC 4584
QY 4750 TGCAGAGGATCCGCTCGCTGSCACCGCCCGCTCGGAGGCTTTTATGTGTAGCGCG 4809
Db 4585 TGCGTGACCGATCCCTCACTGGCGCGGAGGCTCGGCGAGCTACTACTGCTGTCGCGCA 4644
QY 4810 GTGCGCATCTCGGCAACCGCTGACATCGACTGGCAACAGGAGGACCGCTTTCGCGAT 4869
Db 4645 GTTCCGCACTCGGCAACCGCGATATCGACTGGGCGGTTGAAGGTCCGCGCTTCGCGAT 4704
QY 4870 CGAATTTTGTCTTATCTGGAGCAGCATACATGCGCGGATACGTCAGCAATTTAGTGACA 4929
Db 4705 CGCATTTTGTACTATCTGGAACAGCACTATATGCGCGGCTCGGTAGCCAGTTGTCAG 4764
QY 4930 CACGAATGTTTACGCGCTTTGATTTTCGCAACAGCTGCTGATGCCCATCAAGCTCGGG 4989
Db 4765 CATCGCATCTTTCAGCGCTTTGATTTCCGCGATGAGCTGAATGCGTATCAGGGTTTCGGC 4824
QY 4990 TTTTCGCTGGAGCGGATTTTGAACGCAAGCGCTGTTTCCGCGCATAAACCGCATGCC 5049
Db 4825 TTCTCGTGGAGCGGATCTGACGCAAGCGCTGTTTCCGCGCTCACACCGCGATAAA 4884
QY 5050 GATATCAGCAATCTCTATCTGTTGGTGC CGGTACGATCCAGGCGCGGCGTTCGCGC 5109
Db 4885 AATATTAATAATCTCTATCTGTTGCGAGCGCATCTCTGGCGCGGTTATTCAGGC 4944
QY 5110 GTGATCGTTCGCGCAAGGCCACCGCAGGCTGATGCTGGAGGATCGCGCGCAATGAATC 5169
Db 4945 GTAATTTGCTCGGCAAGGCTACCGCAGGATTTGATGCTGGAGGATCTGCGC---TTGAATA 5001
QY 5170 GACAGCTTTTACTTTGAGCAAGTAAACGCAACCATGCGCGTGGCTTCGAAGAGTTTCGCA 5229
Db 5002 GTCCGCTACTCTCTGATCATGCGGTAGACACATGAGGTGGATCGAAAAGCTTTTGCA 5061
QY 5230 CCGCGCGCAAGCTGTTGATGACCGCGCGCGCAGCACGCTGATGCTGATGCTGCT 5289
Db 5062 CCGCGTCAAAACTGTTTGGTCCAAAACCCGACGCGCTGCTGATGCTCTACGCGCTGCT 5121
QY 5290 GTGCTCACTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5349
Db 5122 GCGGTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5181
QY 5350 CCGTGAAGACGCGCAGGCACTATGACGATCTGCAAAATGAAACCCCGCGCGCTTACA 5409
Db 5182 CGTGAATCCGCGCAACAGCGCTGCGCGCAGCTGGAGATGAACCGCTCAGGCTTATG 5241
QY 5410 CCGCGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5469
Db 5242 CCGGATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5301
QY 5470 AGCTGCGCAACAACTGCGCTTTCATCTGGAAGCTTTCGCTATGATGATGATGATGATG 5529
Db 5302 ATATTTGCTGCTTACGCTTTTATCATCTGCGGCGCTTTGCGATGATGATGATGATGATG 5361
QY 5530 AACATTACGCGCTTCGATGACAGCTTTCGCTTACTGCTATCAGCTGCGCGCGCTGCTG 5589
Db 5362 CACGCTACAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5421
QY 5590 GTTTGATGATGCGCGCTTAATGCGCGTGGCGAGAGCGGCTGCTGATCAGCGCTGCG 5649
Db 5422 GCCTGATGATGCGCGATTAATGCGCGTACGCAACCGCACCGCTGATGATGATGATGATG 5481
QY 5650 ATTTAGGACTGCGGCTTCCAGCTCACTAACTGCGCGCATTTGTGAAGATGCGCGAA 5709
Db 5482 ATCTCGTCTGCGGCTTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATG 5541

Qy	5710	ATGGTCTCGCTGATCTCTGCCGAATCTCTGGCTCGAATCAGGCGGGGATTAACGCCGCGATACGC	5769
Db	5542	CGGGACCTCTGCTATCTCTCCCGCTACGCTGGCTGCGTGAAGAGGGGCTTTACCCGAGAGAAATC	5601
Qy	5770	TGACTGCACCGCAACATCGTGCAGCGCTTCGCCTCACTGGCAGCGCGTTTATAGTCGCGGAGG	5829
Db	5602	TCGCCGATCCGCAAAATTCGCACGATTAAGCCGCGTTCGCCCGCGCTGGTGGAAACCGG	5661
Qy	5830	CGGAACCTTATTATCACTTCGCGCGCGATTCGGTTTACCGGGTTTACCGCTTCGCTCGCTCGCCGT	5889
Db	5662	CAGAGCCCTATTATTCGATCGCGCTGGCTTGCCTCGCGGTTTACCGCTGCGTTACAGCGT	5721
Qy	5890	GGGCCATTCGCTACGCTTCGCGCGGCGTTTATTCGCGAAATTTGCGTCAAAGTTTCAGCACGCCG	5949
Db	5722	GGGCGATTGCTACCGCGCAGCAGGTCTACCGCTAAATTCGGTATGAAGTGGTTTCAGCGCG	5781
Qy	5950	GTGTGCACGCTCGGATTTCAGCGCAGCGCACCAAGTAAAGTGAAGTGAAGAACTGGCGCTCTGCG	6009
Db	5782	CTTCACAGGCGTGGGATTCACAGCCAGTTCACACAGCACACAGAGAACTTGGCACTCTCTG	5841
Qy	6010	TGAAGGGGCGAGGTTTGGCCATCACTTCGCGGTGTCTCGTCTCTGAAACCGCTTCGGGCTG	6069
Db	5842	TGGCGGATCGGGTCAGGCGGTTACTTCCGGGTGGCGGTACGCTCCACGCTCCGCTG	5901
Qy	6070	GTCTGTGGCAGCGTCTCTGTTTATTCGTCGCTGACGCTGGCGCAGCGTGGCTTTGCAAG	6129
Db	5902	ATCTCTGGCAGCGCCC-----CGTTTACGCCATGACGTTCTTCGACGCGTTCGCCCTGCAA	5955
Qy	6130	CTTATTACGCGGTGGCGGTAGAGGAAACCAAAACGACACGACGCTTTCACGCCCGCGCAC	6189
Db	5956	TTTTTGACAGCGGGCGCGTAAAGAAAGCCAAACGAGACACAGCCCTCTTTTGGCTCGTAC	6015
Qy	6190	CGCATGATGATCGGTCGCGCCATGTATAGCGCTTAAGATAGCCTTTGCGCGGGATATA	6249
Db	6016	CGCATGATGATCGGTCGCGCCATATAAAGCCTGCGTAAATAACCCCGCGCGGAACATA	6075
Qy	6250	CGCGAAACGGCCAGCGTTTGATGCACACGAGCCATCGTGCAACATGAAGTAGAGCGCGCGTA	6309
Db	6076	CGCGAATGGCCAGCGCTGGTGCAACGAGCCGTATGCACAAATAAATAGNAGCAGGCCGTA	6135
Qy	6310	CGTCGTCAATTCGGGACCAATTCACCTGACGCGGCCACATGCCTTTGACACCGACATAAAT	6369
Db	6136	GAGCGTCAATACCGCGCGGATCCACTGCAAGGGCCAGACACCGGTGTGCCACAGATAAAT	6195
Qy	6370	CAGCACAATCGCAGTACCCCAACACCCCGCATAAAGATCGTTGAGCTTCAAACTTACC	6429
Db	6196	GAGCAAAATCGACAGAGCGCGAACTACGCGCATAGAGTCAATTAACCTTCAAAACAGCC	6255
Qy	6430	GCTGTGGGTTTCATGGTCGACAGATGCCAGCGCCATCCCCAAACCGTGATGATGATTT	6489
Db	6256	TTTTATGGGTTTCATGTGTACAGATGCCAGCCCCCAACCCAGCCATGCATAATGATCTCT	6315
Qy	6490	ATCGACAGCGCCGCTACGATTTTCCATCACCAACCGGTTTGCCAAAGATAAGCACGTT	6549
Db	6316	GTGCGCCAGTGCAGCGGTTATCTCCATTCGGATAACAGTGCACCAACAATCAGGCAATT	6375
Qy	6550	CCATAACGAGCAATTTGTCGTCC	6573
Db	6376	CCAAATCCCAACATAGTTTCTCC	6399

[illegible]

1900 AACTGCCGCTGCTATACGCGGTGGGCGCACTCGCGCCCGCGTTCTCTCGCGCGC 1959
1998 CGTTACCGCTTGTGTTTATCGCTGCGGCTCTGCGGAAACGACGACCGCTCAACGT 1857
1960 TCCATGCGCCCTG---GCCAGCGCTGCGTACGCGGTGTTTATGCTCGCTGGGTACGC 2016
1858 CTTTCATCCGTTATTATCATCTCAGAAACCCCGGATTTTCGCTCGCTGGGACGC 1917
2017 TGCAGGCCATCGCTTCCGCTGTTTCTGCATCTGCGCAGCGGTGCGGCCAGCTGCGGC 2076
1918 TTCAAGGACACCGTTATGCGCTGTTTAAACAGATAGTGAAGCCGTGTGAAGAAATGACG 1977
2077 TATCGTGTGTATCGCCATTTGGGGATTAAACGCCGAAACAGACGATACGCTGAGC 2136
1978 GTACGCTCTGTTAGCCCACTGTGCTCTTAACGACTCTCAGTGTGAAGAGCTGGGC 2037
2137 TCGCTGGCGCGGTGGGTGACGATTTCGTGCATCGACGCGCGAGCCCTACAGCAGCGC 2196
2038 GAAGCGCTCATACAGGTGTGGATTGTCGATCAGTCAGTCAGCGCGGTGTCTCAGGCGC 2097
2197 AGCTGTTTATCACTAATCCGCGTTAAACAGCGCGCTGGAAGCACTTGAATCGGTACGC 2256
2098 AGCTGCGATCAACCAAGCGCGCATGAATACGCTACTGGACGCAATTAATTACCGGACGC 2157
2257 CGATGCTGCGCTGCGGATGCTTTGATCAGCCCGCGGTGGGCGGCATGATGCGC 2316
2158 CCGTTTATAGCTTCGCTGCGCTTTGATCAGCCCGCGGTGCGCTCAGCATCGTTATC 2217
2317 ATGACGTGCTGCGCGCATACGCTTTAGCGGTTCATCACTGGAGCAGCATGCGC 2376
2218 AGCGCATCGCAAGCTGTTCCGCTTTTACACAGCCATGCTTTGGCTGTGATGCGC 2277
2377 AACAGCTGCTGACCGAAGATGTTACGCGCTACGATGTCAGCGATTACGGCGCAGCTGC 2436
2278 GTTCATTGCTGACCAAGTCGATTTACAGCGCATGCGGAAATCCAGACAGCCCTTC 2337
2437 AGCGCGCAGCGGTGCCAGCTGCGCGCATGCTCAGCAGCGGCTGTCAGCAGC 2496
2338 GTTTGCGCAGGGGCAACATGCGCGCTGCGGATATCATTCAGCAGGTTATGTG-CACCGGT 2396
2497 AAGTGTGCTGCGGAGGCGACTGATGCGCAGCAATACGATGTGATTTGCTCGGTGC 2556
2397 CAGCTGTCTTAAGTGGAGCGGCTATGCAACCGCATATGATGTGATTTCTGTGGGCGC 2456
2557 TGGACTGCGAATGCTTGAATGCTGCTGCTGCTCAATTTGACGCCAACAATGAAATG 2616
2457 TGGACTCGGAATGCGCTTATCGCTGCTGCTTTCAGCAGCAGCAACCTGATATGCTAT 2516
2617 CTTGCTGCTGGAGAGGATGCGATCCGCGAGGCAATCATCTGCTGCTTTCATCAGC 2676
2517 TTTGCTTATCGACGCGCAACCCAGCGGGCGGGAATCATAGTGTGTCATTTTCAACACA 2576
2677 CGATCTCAGCGCGGAACAACCTGCTGCTGCAACCGCTGATTACCGTCTGTTGTCAGG 2736
2577 TGAATTTGACTGAGGCAACATCGTTGGATAGCTCCGCTGTTGTTTCATCTAGCCCGCA 2636
2737 TTATCAGGTGCGTTTCTCGCTGCGCCCAATCTGACGCGGGAATTTGTTCCATGCGC 2796
2637 CTATCAGGTACGTTTCCACACGCGCTGTAAGCTGGAACAGCGGCTACTTTTGTATTAC 2696
2797 ATCAGCGGATTTTCCCGCCATCTTTACGCGGGATGGGTGAGGATCTGTGACAAACAC 2856
2697 TTTCTCAGCGTTTTCGCTGAGGTTTTACAGCGACAGTTTGGCCCGCACTTGTGATGGATAC 2756
2857 AGCGGTACAAAGGTAAACCCAGCAGGTGACGCTGCGGGATGGCGGTAACCTTGCTGC 2916
2757 CCGGCTCGCAGAGGTTTAAATGCGGAATCTGTTGCTTGAAGGTGAGGTTATCGGTGC 2816
2917 GCAAGTGTGATTTGATGCTGCGGCTGCGCGCAACCGCACATCTGACGCTGGGTTATCA 2976
2817 CCGCGCGGTGATTTGACGCGCGGTTTATGCGCAATTCAGCAGCTGAGCGTGGCTTCCA 2876
2977 GGTGTTTCTTGACAAAGTGGCAGCTGGCGCAGCGCGCATCGGCTGCGCAGCGCATCTC 3036

2877 GCGCTTTATGCGCAGGAATGCGATGAGCCACCGCATGTTTATCTCTCTCCCATTTAT 2936
3037 GATGATGCGACCGTGCATCAGAAAGCGGTTTATGTTTGTCTACACGCTCGCGCTCAG 3096
2937 CATGATGCGACGGTGCATCAGAAATGTTTATCGTTCTGTGTACAGCTCGCGCTCTC 2996
3097 CGCGCATCGGCTATTGATTGAAGATACCATTAATGTTAACAGCGCGCTGGCGAGAA 3156
2997 GCGGACCAAGATTGTTTAAATTTGAAGACACGCACTATATTGATAATGCGACATTAGATCCTGA 3056
3157 CACCGCTGTGACGACATCGCGCATATGCCAATCAGCAAGCTGCAAGCTGAGTACGCT 3216
3057 ATGCGCGCGGCAAAATATTTGCGACTATGCGCGCAACAGGGTTGCGACTTCAGACACT 3116
3217 GCTGCTGAAAGCAGCAGCATATTACCGATTACCTTGAGCGGCAACATCGATCGATTCTG 3276
3117 GCTGCGAGAAAGACAGGGCGCTTACCATTTACTCTGTGCGGCAATGCCGACGCACTCTG 3176
3277 GCAACAGCAGCGCGGCAAGCGTGCAGCGCTGCGCGCGGCTGTTTTCATGCGCACAC 3336
3177 G---CAGCAGCGCGCTGCGCTGTAGTGAATTTAGTGTGCGCTCTGTTCCATCTACCA 3233
3337 CGGTTACTCTTCCGCTCGCGTGGCTAGCGGATTTGTTAGCAGCGCTGTGCGCCAC 3396
3234 CGGCTATTACCTGCGCTGGCGTTGCGGTGCGCACCGCTGAGTGCACTTGATGTCTT 3293
3397 CGATGCCCTCAGCTCAGCAACATATCGAAGCGTTTGCCTCGTACGAGTGGCGCGCAACA 3456
3294 TAGCTCGGCTCAATTCACCATGCCATTACGCAATTTTGCCTCGGAGCGTGGCAGCA 3353
3457 GCGATTTTTCGCTGTCTGTAACCGCATGCTGTTTGGCGGTAAAGCCGAGCGCTG 3516
3354 GGGCTTTTTCGCACTGCTGAATCGCATGCTGTTTTCAGCGCACCGCGGATTCAGCTG 3413
3517 GCGCTGATGCAACGTTTTCAGCGCTCGATGCGGTTTAAATAGCGCTTTTACGCGG 3576
3414 GCGGTTATGACGCTTTTATGTTTACCTGAAGATTTAAATGCGCGTTTTTATGCGGG 3473
3577 GCAACTGCGCTCGCGCATTAACCGGATTTCTGCGGCAAGCGCGCTGCGCATCGG 3636
3474 AAAACTCAGCTGACGATCGCTAGCTATTTCTGAGCGCAAGCGCTTTCGCTAT 3533
3637 TGAAGCGTGGCGCTGTTGAAATTTCTGCGAAACAGGGAAGAAATAAGAACCACT 3696
3534 AGCAGATTGCAAGCAATTTACGACTCATCTGTTAAAGAGGCACTACATGAACCAACT 3593
3697 TATGTTGTTGGCGAGGCTTTGGCGCTGCGCTGCGGATTCGCTGCAAGCGCGGCG 3756
3594 ACGGTAATTTGGTGGGCTTCGCTGGCTGCGCTGCGCAATTCGTTCAAGCTCGGCG 3653
3757 ATACCAACCACTTACTCAGCAGCGCGCAACAAACCGGCGGAGCGGCTTATGTTGTAG 3816
3654 ATCCCGCTTACTCTGTTGAACAACTGATAAACCGGCGGTTCGGCTTATGTTACGAG 3713
3817 GACAGTGGCTTTCCTTCATGCGCGGACCCAGCGTGAATCACCGATCCGAGCGCATCGAA 3876
3714 GATCAGGGGTTTACCTTTGATCAGGCGCGAGCGTTATCACGATCCGAGTGCATTGAA 3773
3877 GAGTTGTTTCAGCTCGCGGAAATTCGCTCAGCGATTAGTCTGAGCTGATGCGGTAAAG 3936
3774 GAACTGTTGCACTGGCAGGAAACAGTTTAAAGAGTATGTGAACTGCTGCGGTTACG 3833
3937 CCGTCTTATCGCTGCTGCGGAAGATGGCAACAGCTTGAATTAAGCAATTAATCAGCGG 3996
3834 CGGTTTACCGCTGTTGGAGTCAGGGAAGTCTTTAAATTCGATTAACGATCAAC 3893
3997 CTGCTGAGCAGCAGATCGCCACGTTCAATCCGCAAGATGTTAGAGGCTATCGTCAATTT 4056
3894 CGGCTCGAAGCGCAGATTTCAGCAGTTTAAATCCCGCGATGTCGAAGTTTATCGTCA 3953
4057 CTTGCTTATTCAGCTGAAGTATTTAGAGAGGTTTATCTGAACCTCGGCAAGTTCGCTTT 4116

QY 6277 GCCATCGTGCACATGAAGTAGAGCGCGGTACGTGCTCATTCGGGACCAATCCACTG 6336
Db 6165 CCCGTCGTGCACATATAAGTAATAGTAAATACATACGCGGTACATCGGCCAATCCACTG 6224
QY 6337 CAGCGGCCACATCGCTTGCACACCGACATATAATCAGCAATATCGCCAGTACCGCAAAAC 6396
Db 6225 GAGCGGCCACATTCCTGTACTGCCAGATATAATCAGCAGGATCGATATGAGCAAAAC 6284
QY 6397 CACCGGCAATGAATGTTGAGCTCAAACTTACCGGTGCGGTTCATGTTGCGACAGATG 6456
Db 6285 CACGGCAATGAATGTTAACTTCAACCGACCTTACCGCGTTCATGATGTGAAGATG 6344
QY 6457 CAGCGCCATCCCAACCGTGCATGATGTTATATGCGACAGCGCGCTACCAATTCAT 6516
Db 6345 CCATCCCAACCGCGCGTGCATGATGTTATGTCGACGAGCAATCACATTCAT 6404
QY 6517 CACCAACCGTTCGCAACAGATAGACAGCTTCCATACCAAGAGATGTTGTCCTCAT 6576
Db 6405 GCCATCAGGTAACGMAAACGATCAGGCAATCCAAATCCACAAATATTTCTCCGT 6464
QY 6577 TGTGAAAGGGAAGTA 6593
Db 6465 AGAGACGTCGTGCAGCA 6481

RESULT 3
LOCUS I12910 6918 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 13 from patent US 5429939.
ACCESSION I12910
VERSION I12910.1 GI:910887
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6918)
AUTHORS Misawa,N., Kobayashi,K., Nakamura,K. and Yamano,S.
TITLE DNA sequences useful for the synthesis of carotenoids
JOURNAL Patent: US 5429939-A 13 04-JUL-1995;
FEATURES
1. .6918
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 37.5%; Score 2624.6; DB 6; Length 6918;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4007; Conservative 0; Mismatches 2169; Indels 21; Gaps 6;
QY 400 CTTGAACATTTACTGCTCCGGACAGCAAGCGATCGCGTGCCTGCGCGATGCGTCC 459
Db 303 CTTGATCAGTTATTGCCGTGGAGGAGAAACGGGATGTTGTGGTGCCTGCGATGCGTGA 362
QY 460 GGAACGTCGCGCAGGCAAGATGTTCTCTTTTACTGCTGCTGCGCAGCGCGAT 519
Db 363 GTGCGCTGCGCAGCGGAAACGATTTCCGCCCATGTTGCTGTGCTGACCGCGCGCGAT 422
QY 520 ATGGGTGCGAGTACGCAAAATGGCGTTCTCGATCTCGCTGCGCAGTGAATGTTG 579
Db 423 CTGGGTGCGCTGTCAGCATGACGATTTACTGGATTTGGCTTGTGCGTGAATGTC 482
QY 580 CACCGGCAATCGCTGATCTGATGACATTCCTCGATGATGATGATGATGATGATGATG 639
Db 483 CACCGGCTTCGCTGATCTTACGATATGCGCTGATGATGATGATGATGATGATGATG 542
QY 640 GTGCGCTACCGTGCATCGCAATTTGTTGAAAGTGGCGATTCCTCGCGCATTCGCG 699
Db 543 GGAACGCTTACCATTCATTTCTTACGAGAGCATGTGGCAATACTGCGCGGTGCC 602
QY 700 CTGCTTAGCGCGGCAATTTGAAGTGTGTCATTTGACCGCGTTTGTCTGCCATACATAA 759
Db 603 TTGCTGAGTAAGCCCTTTGGCGTAATTTGCGATGAGATGGCTTCAGCGCGCTGGCAAA 662

QY 760 TCTGAGCGATTTGCTGAACCTCTCCGCTCGCGCTGACGGCTTAGTGCAAGGGCAA 819
Db 663 AATCGCGGGTTTCTGAACTGTCAAAACGCTATCGGCATGCAAGGATTTGGTTCAAGGTCAG 722
QY 820 TTCCAGGATCTGACAGACGCGACAGCGAGCGCGAGCGCGGAGGAGCGATCGCCATGACCAAC 879
Db 723 TTCAAGGATCTGCTGAAGGGGATGAAGCGCGAGCGCTGAAGCTATTTTGTATGACGAAT 782
QY 880 GAATGAAACACAGCGTCTGTTTCGCGCCACGCTGCAAAATGCGCGCGGATTCGCGCTGAC 939
Db 783 CACTTTAAACACAGCACGCTGTTTGTGCTCTCATGACAGATGGCCTCGATTTGTCGAAT 842
QY 940 GCTTACCGCAGGTGCGCAAGACTTAGCTTCTTCGCCCAGGATTTGGCGCCAGCGTTT 999
Db 843 GCCTCCAGCGAAGCGCGTGAATGCTGCTGATGCTTTTCACTTGAATCTTTGGTCAGGCAATTT 902
QY 1000 CAACTGCTCGACACCTCGCGACGCTTGAACACACACCGGTAAAGATGTGCAACAGGAT 1059
Db 903 CAACTGCTGACGATTTGACCGATGCGCATGACCGACACCGGTAAAGATGACGAATCAGGAC 962
QY 1060 CAGGCAAAATCCACGCTGCTGACAGATGCTGCGTGTGACGCGCGGAAAGCTTCGCTGCGC 1119
Db 963 GCGGTAAATCGACGCTGCTCAATCTGTTAGCGCGAGCGGCTTGAAGACGCTCTGAGA 1022
QY 1120 GATCACCTGCGCAGCGACAGATGCACACCTTGCCTGCGCTGCCATCGCGGATCGCCACT 1179
Db 1023 CAACTTCTGACGTTGCGCAGTGAATCTCTTCGCGGCTGCAACACCGGACGCGCACT 1082
QY 1180 GCGCAATATATGACGCGCTGTTTAAATCAACAGCTAGCGATATTAACACTGAGCGCGGCTC 1239
Db 1083 CAACTATTTATTCAGCGCTGTTTGAACAAAACTTCGCTGCGCTCAGTTAAGATGCTGC 1142
QY 1240 AGCGGTGCGCACCTTTGCGGTGATGCGCGCGCTCTACAGCCACTTTTCAAGCGGTGCG 1299
Db 1143 A-----TGAGCCATTTTCGCGCGATCGCACGCGCTTTTACAGCCATGTTTCGCGCAATAC 1197
QY 1300 AGCGGTGACCAAAAGCTGCTGCGCGCGCGCGCGCATCGCATCATTCATCATCGACGCGG 1359
Db 1198 AGAATCTCGCTCAGGAACCTGCTGCGCGCGGTCATCGGCTGACCTTTTATTCAGCAATACG 1257
QY 1360 ATGCGCGCACCTTTGCTTAGCGACGAAACGATCGATTTTGTGCGCTGCGCAACAGACGC 1419
Db 1258 ATATTAAACATTTGATCGATAGCAAAACATTTGATTTTCATTCGCTCGGACAGACGCG 1317
QY 1420 ATCTGCGCGTTGCTGCGCGCGCTGTTGCTGCTGCTGCTGCGCGCGCGCTGCTGCTGCG 1479
Db 1318 ATCCCCCGCGCGTTAAACGCGCGTGTACACCTGCGCGCTCATCTCTGCGCGCGCTCAA 1377
QY 1480 TGTTTCGCGTGTATCGACGATCTCGCGCTCTGACCGATATGCTGCGCGCAACTGCGCTG 1539
Db 1378 TGCTGAAGCTCATCAATGAAATGGCGCGCACCGCATGCTGCTGCGCGCAACTCCCC 1437
QY 1540 CGGTACTGAAAGCATTTGAACATCGATGCGTATCGCGACGAAATGGAAGCGCGCGCG 1599
Db 1438 AGGCATTTAAGATCTGCGCGCTGATGCGCTCATTTGTTGATCAATGAAACCGCGCGCG 1497
QY 1600 GATTGCTGCTGAAGCGCTGCAATCTGCGCTTTGTTTCGCTGCGCTTTCGCGCTTTCGCGGTCA 1659
Db 1498 CGCTGTTGCTGAAGCATGCGGACTGCGCTTTATCTGCTGCGCTGCGCGCTGCTCTCA 1557
QY 1660 ATCTGGAAGCGGATTCGCTTGGGTGATGCGCTTCCGTTTTCGACAGGATGACAAAG 1719
Db 1558 ATCTGGAACCGGATATGCGCTGCGGTTATGCGCTTTTGAATACGGGACAGCGCGCG 1617
QY 1720 CGCTGAAACGTTTTCAGGCGCAGCAGCATATCTATGATCGCATATGCTGCTGCTGCGCG 1779
Db 1618 CTCGGAACGTTATCGCGCGCTGAAATTTTATGATGCGCTATGCTGCTGCTGCTGCTG 1677
QY 1780 AGTGATCTTCAACACAGCGCGCGGCTTTAATTTGAACGAGCGCGCGGATTTACATCAT 1839
Db 1678 GTGCTATTTGCGGAACACAGCCACAGATGGGCTTAGCGCGCTTAGCGCGCGGCAAAAGCTTCA 1737
QY 1840 GCCTGTCGCGCTGCGCAAAATCAGCCAGATGCTGCGCGCTTTGATTTTCCACGTCAGC 1899

1738 GTTCTTTCGCACTGGCGCAATCAGCAGCTTGTCTCTGAACTGGATTTTCCTCCGCAAG 1797
1900 AACTGCGCGCTGCTATACGCGCGTGGGCGCACTCGCGCCCGCGTTCCTCTGCGCGC 1959
1798 CGTTACCGGTGTGTTTCTATCGCGTGGGCGCTCTGCGGAAAGCAACGACCGTCAACGT 1857
1960 TCATGCGCGCTG- --GCCAGCGCTGCGTCAGCGCGTGTGTTATGCTCGCTGGGTACGC 2016
1858 CTTTCATCCCGTTTATTTATCATCTCAGAAAAACCCCGGATTTTCGCTCGCTGGGACGC 1917
2017 TGCAGGCGCATCGCTTCGCGCTGTCTGATCTGCGCAGGCGTGGCGGCGTGGCGC 2076
1918 TTACGGGACACCGTTATGCGCTGTGTTTAAACAGATAGTGAAGCGCTGTGAAGAAATGACG 1977
2077 TATCGCTGTGTATCGCGCATTTGGGGGATTAACCGCCGAAACAGACGCACTAGCTGGAGC 2136
1978 GTACGCTCTGTGTAGCCCACTGTGCTGCTTACGACCTCTCAGTGTGAAGAGCTTGGCGC 2037
2137 TCGCTGGCGCGGTGGGTGACCGGATTTCTGATCAGCGCGCAGCCCTACAGCAGCGC 2196
2038 GAAGCGCTCATACACAGGTGGTGGATTTTGGCCGATCAGTCAGCGCGCTCTCAGGCGC 2097
2197 AGCTGTTTATCACTCATGCTCGCGGTAAACAGCGCGCTGGAAGCACTGGAATCGGTACGC 2256
2098 AGCTGGCGATACCCACAGCGCGCATGAATACGCTTACGACGCGATTAATTAACCGAGCGC 2157
2257 CGATGCTGCGCTGCGCATGTGCTTTGATCAGCGCGGTGGCGCGCATTTAGTGGC 2316
2158 CCGTTTACGCTTCGCTGGCTTTGATCAGCGCGGTGCGCTGCGTACGATCGTTTATC 2217
2317 ATGAGTTGGTTCGCGCGCATCACGCTTTAGCGCTTTTCATCAACTGGAGCAGCATCTGC 2376
2218 ACGGATCGGCAAGCTGCTTCCCGCTTTACCAACAGCCATGCTTTGGCTGCTCAGATGC 2277
2377 AACAGCTGTGACCGACGATCGTTAGCGCTAGGATGTCAGGATTCAGGGCGAGCTGC 2436
2278 GTTCATGTGACCAACGCTGACCTTTCAGCAGCGATGCGGAAATCCAGACAGCGCTTC 2337
2437 AGCGCGCAGCGGTTCAGCGCTGCGCGCATCGTCAGCAGCGCTGTCAGCAGCAGC 2496
2338 GTTTGGCAGGGGCACTGCGCGCTGCGATATCATTTAGCAGAGTTATGTG- CACCGT 2396
2497 AAGTCGTGCTGCGGAGGCGCACTGATGCGCAGCAATACGATGATGATTTGGTTCGCTGC 2556
2397 CAGCCTGTCTTAAGTGGGAGCGGCTATGCAACCGCATTTATGATCTGATCTCGTGGGCGC 2456
2557 TGGACTGGGATGCTGATTTGCGCTGCTGCTGCTCAATTTGACAGCCACACTGAATG 2616
2457 TGGACTCGGAATGGCTTTATCGCCCTGCGTCTTACGACGAGCAACCTGATATCGTAT 2516
2617 CCTGTTGCTGAGAGCGATGCGCATCCGCGAGGCAATCATCTGCTGCTTTCATCACAG 2676
2517 TTTGCTTATCGACGCGCACCCAGCGGGGCGGAAATCATAGTGTGTCATTTACCACGA 2576
2677 CGATCTCAGCGCGCAACAACTTTCGTGCTGCAACCGCTGATTAACCGCTGCTGCTCAGG 2736
2577 TGATTTGATGAGAGCAACATCTGTTGATAGCTCGCTGCTGCTGCTGCTGCTGCTGCTG 2636
2737 TTATCAGTGTGCTTTCGCTGCGCGCATCTGGAACGGGATTTATGTTCCATCGC 2796
2637 CTATCAGGTACGCTTTCCACAGCGCGCTGTAAGCTGAACTGAACGCGCTACTTTTGTATAC 2696
2797 ATCAGCGATTTTTCGCGCATCTTTACGCGCGATGGGTGACGATCTGTGGACAAACAC 2856
2697 TTCTCAGCGTTTCGTGAGTTTACAGCGACAGTTTGGCCCGCATCTGTGGATGGATAC 2756
2857 AGCCTCAACAGGTAAACCAACGAGGTGACGCTGCGGATGCGCGTGAACCTTGTGCTGC 2916
2757 CCGCTGCGCAGAGGTAAATGCGGAAATCTGCTTGGTTGAAAGGGGTACAGTTATCGGTGC 2816
2917 GCAAGTGTGTATGATGTTCGCGGCTTCAGCGCGACGCGCATCTGACAGCTGGGTATCA 2976

2817 CCGCGCGGTGATTTGACGGCGCGGGTTATGCGGCAAAATTCAGCACTGAGCGTGGCTTCCA 2876
2977 GGTGTTTCTTGACAAAGATGCGAGCTGGCGCAGCGCGCACGCCCTGCAAGCAGCCATCCT 3036
2877 GCGTGTATTTGCGCAGGAATGCGATTTGAGCCACCCGATGGTTTATCGTCTCCCATTTAT 2936
3037 GATGATGCCACCGCTGATCAGCAAGCGGTTATCGTTTGTGTTCTACAGCTGCGCTCAG 3096
2937 CATGATGCCACCGCTGATCAGCAAAATGTTATCGCTTCGTGTACAGCTGCGCTCTC 2996
3097 CGCCGATCGGCTATTGATTTGAATGATACCCATTAACGTTAACAGCCCGCGCTGGCGGAGAA 3156
2997 GCGGACCAAGATTTGTTAAATTTGAAGACACGCACTATATTGATTAATGAGATAGATCTCA 3056
3157 CACCGCTGTCAGCACATCGCGCATGCGCAATGCGCAATCAGCAAGCTGCGAGCTGAGTACGCT 3216
3057 ATGCGCGCGCAAAATTTTGGACTATGCGCGCAACAGGTTTGGAGCTTTCAGACACT 3116
3217 GCTGCTGAAGACGACGCGCATATTACCGATTTACCTGAGCGGCAACATCGATTCGATTCG 3276
3117 GCTGCGAAGAAACAGGCGCGCTTACCCATTACTCTGTCGGGCAATGCGAGCATTCG 3176
3277 GCAACAGCAGCGCGCGCAAGCGCTGAGCGCGCTGCGCGCGCGCTGTTTCATGCCACAC 3336
3177 G---CAGCAGCGCGCGCTGAGTGTGATTTAGCTCCGCTGCTGTTCCATCTCTACAC 3233
3337 CGGTTACTCTTTCGCTCGCGCTGCGCTGAGCGAGTTGTTAGCAGCGCTGTTGCGCCAC 3396
3234 CGGCTATTACTGCGCTGCGGTTGCGTGGCGCAACCGCTGAGTGCATTTGATGCTT 3293
3397 CGATGCCCTCAGCTCAGCCAAACATATCGAACGCTTTTCCCGCTCAGCAGTGGCGGAAACA 3456
3294 TAGCTGCGCTCAATTCACCATGCAATTCAGCATTTTGCCTCGAGCGCTGCGAGCAGCA 3353
3457 CGGATTTTTCGCTGCTGTAACCGCATGCTGTTTGGCGCGTGAAGCGAGCAGCGCTG 3516
3354 GGGCTTTTTCGCTGCTGTAATTCGATGCTGTTTATGCGGAGCCCGCGGATTCACGCTG 3413
3517 CCGGTGATGACAAAGTTTTCGCGCTGATGCGGTTTAAATAGCGCTTTTACGCGG 3576
3414 CCGGTTATGAGCGTTTATGTTTACCTGAAAGATTTAAATGCGCTTTTATGCGGG 3473
3577 GCAACTGCGCTGCGCGATATAACCGGATTTCTGCGGCAAGCGCGCTGCGCATCGG 3636
3474 AAACTCAGCTAGCCGATCGCTACGATTTCTGAGCGGCAAGCGCTTTCGCTGAT 3533
3637 TGAAGCGCTGCGCGCTGTTGATTTCTGTCGAACCGAGGAAAGAAATAACGAGCACT 3696
3534 AGCAGCATTTCAAGCCATTTATGACGACTCATCTGTTAAAGAGCGACTACATGAAACCAACT 3593
3697 TATGTTGCGCAGCGCTTTGCGCGCTGCGCTGCGGATTCGCTGCAAGCGCGCGG 3756
3594 ACGGTTATGTTGCGAGCTTCGCTGCGCTGCGCATTTGCAATTTCTTACAGCTTGGGG 3653
3757 ATACCAACCACTTACTGAGCAGCGCAAAAACCGGCGGACGCGCTATGTTGTTGAG 3816
3654 ATCCCGCTTACTGCTTGAACAACGATATAACCGCGCGTTCGGCTTATGTTCTACGAG 3713
3817 GACAGTGGCTTACTTTCGATGCGGACCCACGCTGATCACCGATCCAGCGCATCGAA 3876
3714 GATCAGGGTTTACTTTGATGCGAGCGCGGCTTATCACCCGATCCAGGTGCCATGAA 3773
3877 GAGTTGTTTACGCTGGCAGGAAATTCGCTCAGCGATTTACGTCGAGCTGATGCCGGTAAACG 3936
3774 GAATGTTTGCATGCGCAGGAAACAGTTAAAGAGTATGTCGAACCTGCTGCGGTTACG 3833
3937 CCCTTCTATCCCTGCTGCGGAAAGATGGCAACAGCTTGTGATACGAAATAATACGCG 3996
3834 CCGTTCACCGCTGCTTGGGAGTCAGGGAAGGCTTTTAAATTAACGATACGATCAAAAC 3893
3997 CTGCTGAGCAGCAGATCGCCACGTTCAATCCGCAAGATGTAGAGGCTATCGTCAATTT 4056
3894 CCGCTCGAAGCGCAGATTCAGCAGTTTAAATCCCGCGATGTGCAAGGTTTATCGTCAGTTT 3953

6105 CAACCGTTTGGAGTAGCCCTTGGTGAATATAGCGAATGCCAACGTTGATGCACCAG 6164
6277 GCATCGTCACCATGAAGTAGAGCGCGCGTACGTCGTCATTCGCGCACCAATCCACTG 6336
6165 CCCGTCGTCACCATGAATAGAGTAATCATACGCCGTCATACCTGCGCCAATCCACTG 6224
6337 CAGCGGCCCATCGCTTGCACACCGACATATAATCAGACAAATCGCCAGTACCGCAAAAC 6396
6225 GAGCGGCCCATCTCTGTACTGCCAGATAAATCAGCAGGATCGATATGACGCAAAAC 6284
6397 CACCGCATAAAGATCGTTGAGCTCAAACTTACCGCTGTCGGTTTCATGTCGACAGATG 6456
6285 CACCGCATAAAGATCGTTAACTTCAACCGCCTTTACCGCGTTTCATGATGTGAAGATG 6344
6457 CAGCGCCCATCCCAACCGTCATGATGATTATTCGACAGCGCGCTACCATTTCCAT 6516
6345 CCATCCCAACCGCGCTGATGATGATTGTCGTCAGTCAGCAATCATTCCAT 6404
6517 CACCACCGCTTGCACCAAGATAGCAGCTTCCATACACGAGCATTTGTTCCGTCAT 6576
6405 GCCATCACGTAACGAAACGATCAGGCGATTCBAATCCACACATAATTTCTCCGCT 6464
6577 TGTGGAAGGGAAGTA 6593
6465 AGAGACGCTCGGCAGCA 6481

RESULT 4
AY166713
LOCUS
DEFINITION
Pantoea stewartii zeaxanthin-digluco-2002
complete sequence.
AY166713
AY166713.1 GI:27228289
Pantoea stewartii
Pantoea stewartii
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pantoea.
deSouza,M.L., Kollmann,S.R. and Schroeder,W.A.
Carotenoid Biosynthesis (WO 02/079395 A2)
Unpublished
deSouza,M.L., Kollmann,S.R. and Schroeder,W.A.
Direct Submission
Submitted (18-OCT-2002) Biotechnology Development Center, Cargill
Inc., P.O. Box 5702, Minneapolis, MN 55440, USA
Location/Qualifiers
1. .6586
/organism="Pantoea stewartii"
/mol_type="genomic DNA"
/strain="ATCC 8200"
/db_xref="ATCC:8200"
/db_xref="taxon:66269"
169. .6399
/gene="zeaxanthin-digluco-2002 biosynthetic operon"
169. .1080
/gene="crtE"
169. .1080
/gene="crtE"
/codon_start=1
/transl_table=11
/product="geranylgeranyl pyrophosphate synthase"
/protein_id="AA85596.1"
/db_xref="GI:27228290"

1090. .2385
/gene="crtX"
1090. .2385
/gene="crtX"
/codon_start=1
/transl_table=11
/product="zeaxanthin glucosyl transferase"
/protein_id="AA85597.1"
/db_xref="GI:27228291"
/translation="MSHPAVIAPPPFSSHVRALONLAQELVARGHRTVFFQOHDCKALV
TGSIGQTVGLQTHPPGSLSHLLHAHPLGSPMLRLINEMAPTSDMLCRLPAAFH
ALQEGVTVDMQEPAGAVVAASGLPFFVSACALPLNNRPGPLPAWFEFFGTSDAAR
ERYTSEKIYDMRHRDRIAHACRMGLAPREKHLHCFSPLAQISOLILPELDPK
ALPCCFHAVGLRQPGTSGSTSYFSPDPKPIFASLGLQGHRYGLFRTAKACEE
VDAOLLAHCGSLGATAGELARGDIQVDFADQSAALSOAQITITGGMNTVLDAI
ASRPLPLAFLAFOGPGVASRIVVHGIGKRSRTTSHALARQIRSLTLNTDYPQMT
KLAARLPLAGTTPAAADIVEQAMTCQPLVSGQDYATL"
2369. .3517
/gene="crtY"
2369. .3517
/gene="crtY"
/codon_start=1
/transl_table=11
/product="lycopene beta-cyclase"
/protein_id="AA85598.1"
/db_xref="GI:27228292"
/translation="MQPHYDLILVAGLANGLIALRLOOQHPDMRILIEAGPEAGN
HTWSPBEDLTNQHRTIAPLVVHWPDYQVRFQRRHVNSSGYCVTSRHFAGILRQ
QFGHLMHTAVSAVHAESVOLADGRIIHASTDIGRGYTPDSALRVGFOAFIGQWQ
LSAHLGSLPIIMDATVDQNGYRFVTLPLSATALLEDTHYIDKANLOARARONI
RDYAAQWQPIQTLREEGALPITLTGDNRFQWQFQWQPCAGSLRAGLPHFTTGYSLP
LAVALADLSALDVTSSVHOTIAHFAQWQWQGFRLNRLMLFLAGPAESWRVM
QRFYGLPEDLIARFYAGKLTVDRLILSLGKPPVPVFAALQAIMTTHR"
3529. .5007
/gene="crtI"
3529. .5007
/gene="crtI"
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="AA85599.1"
/db_xref="GI:27228293"
/translation="WKPTTVIAGFGGLALIRLOQAGIPVLLLEQRPKPGRAVYQ
EQGTFDAGPTVITDPSAIEELFALAGKQKQVLLPVTFFYRICWESGKVFNDND
QAQLEAQIQFNPRDVAGYAFIDYSRAVFNKGLKGTFFLSPKMLRAAPQALAKL
QAMRSVSKVAGYIEDEHLQAFSLVGNPNFATSIYTLIHALERWGWPPRG
GTGALVNGIKLFODLGEVVLNARVSHMETVGDKIQAQVLEDGRFRFETCVASNADV
VHTVDDLISOHPAAKQAKLQSKMSNSLVLYFGLNHHDDQAHHTVCPGPRVREL
IHEIFNHDGLAEDFSLYLHAPCVTDPSLAPGCGSYVVLAPVHLGTANLDWAVEGPR
LRDRIFDLYEQHYMPLGRSLQVTHRMFTTDFRDELNAWQGSFVSVEFILTQSAWFRP
HNRDKHLDNLYLVGAGTHPGAGIPGVIGSAKATAGLMLEDLI"
5043. .5933
/gene="crtB"
5043. .5933
/gene="crtB"
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="AA85600.1"
/db_xref="GI:27228294"
/translation="MAVSKSFATASTLFDKTRPSVLMVYACWCHCDVDDOTLGF
HADQSPQEPQRLQOLEMKTRQAYAGSQMHEPAPAFQEVMAHDIAPAYAFDHLG
FAMDVRETRYLTLDLTRYVHGVGLMMAQINGVDNATLDRCLAGLAFQITNI
ARDIVDDAQVGRCYLPESWLEBGLTKAYAAPENRQALSIRAGLREAEPPYVSSM
AGLAQLPLRSAAWATATQVYRKVQKVEAQKQAWDRHQSTSTAEKLTLLLTASGQA
VTSRMKTYPPRPAHLWQRP"
complement(5872. .6399)
/gene="crtZ"
complement(5872. .6399)
/gene="crtZ"
/codon_start=1
/transl_table=11
/product="beta-carotene hydroxylase"

gene
CDS
gene
CDS
gene
CDS
gene
CDS

/protein_id="AA085601.1"
/db_xref="GI:2728295"
/translation="MLMIAWNLIVFVVVMEVVAALAHKVIHMGWNGWHLSHHEPR
KGFVNLDYAVVFAIVSIALIYFGSTGIWPLQWAGMTAYGLLYFVWHDGLVHQWR
PFYTPRKGYLKRLYLMAHMHAVRKGCVSFGFLYAPLSKLOATLRERHARSQA
ARDEQDGVDTSSSGK"

ORIGIN

Query Match 37.4%; Score 2620; DB 1; Length 6586;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4016; Conservative 0; Mismatches 2170; Indels 24; Gaps 7;
QY 373 CAGCTGCAGCAAAATTTCAGCGCATCTTGAACGATTTACTCCCTGCCGACAGCAAGC 432
DB 220 CAGTTGCTGGCTGATCATAGCCGCTTGATCAGTTACTCTCCCGTTCCAGGGTAGCGG 279
QY 433 GATCGCTGCTGCCGCGATCGCTGCCGAAACGCTGGCGCAGCGCAAAACGTTATTCGTCT 492
DB 280 GATTGTGGTGCCGCGATGCTGAGGACGCTGGCACCGGCGCAACGTTATTCGTCTCG 339
QY 493 TTATTACTGCTGCTGGCAGCGCGGATATGGGTGGAGCTGACGCAAAATGGCGTTCTC 552
DB 340 ATGCTGCTGTTATTAAACAGCGCGCATCTTGCTGTGCGATCAGTCAAGGGGATTACTG 399
QY 553 GATCTGCTGTGCTGAGTGAATGTGCACGGCGATCGCTGATTTCTGATGACATTCCTC 612
DB 400 GATTTAGCTGCGGCTTGAAATGTGATGCTGCTGCTGATTTCTGATGATATGCC 459
QY 613 TCGATGGATTAACCGCAGATGGTGTGTCGCCCTTACCGTGCATCGCGAATTTGGTGAA 672
DB 460 TGCATGGACGATGCGCAGATGGTGGGGCGTCCCAACATTCACACGATGACGGTGA 519
QY 673 AACGTGGGATTTCTCGCCGCATTCGCGTGTCTAGCCGCGCAATTTGAAGTGAATGCCA 732
DB 520 CATGTGGCGATTCGCGCGCTGCTTTACTCAGCAAGCGTTTGGGGTGATTCGCGAG 579
QY 733 GCACCGGTTTCCCTGGCCATACATAATCTGAAGCGATGTGTAATCTCCGCTCGCTC 792
DB 580 GCTGAAGGTCGACGCGGATGCCAAAACCTCGCGCGGTGTCGAGCTGTCCACTCGATT 639
QY 793 GCGCTGCAGGGCTTAGTCAAGGCAATTCAGAGATCTGCAGACGCGACGACAGCGCG 852
DB 640 GGCATGCGGGTCTGGTTACGGCCAGTTTAAAGACCTCTCGAAGGCGATAAACCCGCG 699
QY 853 AGCCCGGAAGCGATGCCATGACCAACGAACTGAAACACGCGTCTGTTTCGCCCAAG 912
DB 700 AGCGCGGATGCCATCTGCTAACCAATCAGTTTAAACCCAGCACGCTGTTTTCGCGTCA 759
QY 913 CTGCAATGGCGCGATTCGCGCTGACGCTTACCGCAGGTGCGCAAGACTTAGCTTC 972
DB 760 ACGCAATGGCGTCCATTGCGGCCAACGCGTCTCGAAGCGCGTGAGAACCTGCATCGT 819
QY 973 TTCCGCCAGGATTTGGGCGAGCGTTTCACTGCTCGACGCTCGCGAGGTTGCAAA 1032
DB 820 TTCTCGCTCGATCTCGGCCAGGCTTTGAGTTGCTTGAGATCTTACCGATGGCATGACC 879
QY 1033 CACACCGGTAAAGATGTGCACAGGATCAGGGCAAAATCCACGCTGGTACAGATGCTCGGT 1092
DB 880 GATACCGGCAAGACATCAATCAGGATGCAAGTAAATCAACGCTGCTCAATTTATTAGGC 939
QY 1093 GCTGAGCGCGGAAGCTGCTGCGGATCACTGCGCAGCGGAGATGACACACTTGGC 1152
DB 940 TCAGCGCGGTGGAAGACGCTGCGACAGCATTTGCGCTGGCCAGTGAACACCTTTCC 999
QY 1153 TCGCCTGCTCATCGCGGATGCCACTCGCCAAA---TATATGACGCGCTGTTTAAATCAA 1209
DB 1000 GCGGATGCAAAACGGCAATTCACACCCAACTTTTATTTCAGGCGCTGGTTGACAAA 1059
QY 1210 CAGCTAGCGATATTCAACTGAGCGGCTCAGCCGCGTGGCCACTTTTCGGGTGATCGGCG 1269
DB 1060 AACTCGCTCGCGTCAAGTTAAGATGCTGCA-----TGAGCCATTTTGGGGTATCGCAC 1114
QY 1270 CGCCGCTCTACAGCCACTTTACGGGTGACGGGCTTAGCAAAACGCTGCTGGCGCGG 1329

DB 1115 GCCTCTTTTTCAGCCATGTTTCGGCTCTGCAAAACCTTGTCTCAGGAATTAGTGGCCGCG 1174
QY 1330 GCCATCGCATCACATTCATCCAGCAAGCCGATGCCCGCACTTTGCTTAGCGACGAACGCA 1389
DB 1175 GTCATCGTTAGCTTTTTCAGCAACATGACTGCAAAAGCGCTGCTAAACGGGACGATA 1234
QY 1390 TCGATTGTTGCTGCGGCGCAACGACGCAATCTGCGGTTTCGTCGGGTCGTCGGCCCGT 1449
DB 1235 TCGGATTCAGACCGTTCGGACTGCAAAACGCAATCTCCCGGTTCTTATTCGCACTGCTG 1294
QY 1450 ATCGCTGCTGCTGCGCGGCGCTGCTGCTGTTTTCGCTGATCAGCATCTCTCGCTCCT 1509
DB 1295 ACCTGGCGCGCACCCACTCGAACCTCGATGTTTACGACTGATCAATGAATGGCACGTA 1354
QY 1510 GCACGATATGCTGTGCGCGCAATGCTGCGGTACTGAAAGCATTTGAACATTCGATGGCG 1569
DB 1355 CCAGCGATATGCTTTGCGCGGAACTGCGCGCGCTTTTCATGCGTTCGATAGAGGCG 1414
QY 1570 TGATCGCGACGAATGGAAGCGGCGGCGATTTGGTCTGTAAGCGCTGCAATCTGCGCT 1629
DB 1415 TGATCGTTGATCAAAATGGAAGCGGCGAGTGCAGTAGTCGCAAGAGCTCAGGCTCTGCC 1474
QY 1630 TTGTTTCGCTGCGCTGCGCTTTCGCGTCAATCGTGAAGCGGATTCGCTTTCGCGTGA 1689
DB 1475 TTGTTTCGCTGCGCTGCGCTTTCGCGTCAACCGCGAAACCGGTTTTCCTCTGCGCGTGA 1534
QY 1690 TGCCCTTCGCTTTTGCACAGGATGCAAAAGCGCTGAAACGTTTTCAGGCGCAGCAGATA 1749
DB 1535 TGCTTTTCGAGTACGGCACCGCATGCGGCTCGGAAACGCTATACACACAGCAAGAAA 1594
QY 1750 TCTATGATCGCATCATGCTGCTGTCGCGGACGCTGATCCTCAACACAGCGGCGGCTTTA 1809
DB 1595 TTTATGATGCTGCTGATGCGAGCTCAGCATCGTGTGATCGCGCATCATGATCAGAAATGG 1654
QY 1810 ATTTGACGAGCGGCGCGGATTAATCAGTGCCTGCTGCGCGCTGCGCAAAATCAGCCAGA 1869
DB 1655 GTTTAGCCCCGCTGAAAACTGCAATCATTTTCTTCTCACTGGCAAAATCAGCCAGT 1714
QY 1870 TGCTGCGCGCTTTGATTTTCCAGTCAAGCAATTCGCGCGCTGCTTATCAGCGCTGGGCG 1929
DB 1715 TGATCCCCGAACTGATTTTCCCCGCAAGCGCTGCGACAGCTGCTTTTCATGCGGTTGAC 1774
QY 1930 CACTCGCG---CCCGGTTTCTCTGCGCGCTCCATGCGCGCTGCGCAGCGCTGCTGTC 1986
DB 1775 GCTTACGCAACCCCGAGGACGCGCGGCTCATCAACTTCTTATTTTCGCTCCCGGACA 1834
QY 1987 AGCGGTGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046
DB 1835 AACCCCGTATTTTTCCTGCTGCGCACCTTGCAGGAGCATCGTTATGCGCTGTTTTCAGGA 1894
QY 2047 ATCTGGCGCAGCGCTGCGCGCATGCTGCGCTATGCTGCTGATGCCCATTTGTGGGGAT 2106
DB 1895 CCATCGCAAGCGCTGCAAGAGGTGGATGCGAGTTACTGTTGCGCACACTGTGGCGGCC 1954
QY 2107 TAAACGCGCAACAGACGATCAGCTGAGTCTGCTGCGCGCGCTGCGTGGTGAAGATTGCG 2166
DB 1955 TCTAGCCACGACGAGGAGTGAATGCGCGCGGCGGGAATTCAGGCTTGTGGATTTTG 2014
QY 2167 TCGATCAGCGCGCAGCCCTACAGCACGCGAGCTGTTTATCACTCATCGCGGTTAAACA 2226
DB 2015 CCGATCAATCCGACGACTTTTTCAGAGGCAAGTTGAAATCACATCGTGGGATGAATA 2074
QY 2227 GCGCGCTGGAAGCACTGGAATGCGGTAGCGGTGCTGCGCTGCGGATTTGTTTGTATC 2286
DB 2075 CGGTACTGGAAGCTTATGTTCTCCGCAACACGCTACTGCGCGCTGCGCTGCGATTTGATC 2134
QY 2287 AGCCCGGCTGCGCGCGCATTTGATGCGCATGAGCTTGGTGGCGCGCATCACGCTTTA 2346
DB 2135 AACCTGGGCTGGCATCAAGATTTGTTTATCATGGCATCGGCAAGCGTGGCTCTCGGTTA 2194
QY 2347 GCGGTGTTTCACTCACTGAGCAGCATCTGCAACAGCTGCTGACCGACCATCGTTACGCG 2406

QY 4567 ACATCGCTGAAGCGTAAAGCGATGAGCACTCGCTGTTTGTACTCTATTTTGGCCCTGAAT 4626
DB 4411 AAAAAATCCAAATCCAAAGCGTATGATGAATCACTCACTGTTTGTACTCTATTTTGGCTCAAC 4470
QY 4627 CAGCGCGATGAACAGCTCGCGCACACACCGTCTGTTTGGGCCCGCTTATCGTCAGTTG 4686
DB 4471 CATCATCAGATCACTCGCCCATCATACCGTCTGTTTGGGCCACGCTACCGTGAATG 4530
QY 4687 ATCGATAGATTTTCAACAGCAGCAGCTGGCAGACGATTTTTCATCTTTCACTGACCGG 4746
DB 4531 ATTCAAGAAATTTTAAACCATGATGCTGGCTGAGGATTTTTCGCTTTATTTTACAGCA 4590
QY 4747 CCTCGCAGCAGGATCCGCTCGCTGTCACCGCCCGCTGGCGAGCTTTTATGTTAGG 4806
DB 4591 CCTGTGTACGGATCCGCTCACTGGCAGCGGAGGTGGCGAGCTATTTATGCTGGCG 4650
QY 4807 CCGGTGCGCATCTCGGCACCGCTGACATCGACTGGCAACAGAGGACCGCTTGGCG 4866
DB 4651 CTGTGTTCCACATTAGGCACGGGAACTCTGACTGGCGGTAGAGGACCCGCACTGGCG 4710
QY 4867 GATCGAAATTTTGTCTTATCTGAGCAGCACTACATCGCGGATTAAGTCAGCAATTAAGT 4926
DB 4711 GATCGTATTTTGTACTTACCTTGAGCAACATTAATACATGCTGGCTTGGCAAGCCAGTTGGT 4770
QY 4927 ACACACAGATTTTACGCGCTTGTATTTTCGGACACGCTGCATGCCATCACGCTCG 4986
DB 4771 ACGCACCGTATGTTTACGCGCTTGTATTTTCGGACGAGCTCAATGCTGGCAAGGTTG 4830
QY 4987 GGTGTTTCGCTGAGCGGATTTTGAAGCAAGCGCTGTTCCGCCGATCAACCGCAT 5046
DB 4831 GCTTCTCGTTGAACTTATCTGACCCAGAGCGCTGTTCCGACCAATCAACCGCGAT 4890
QY 5047 GCCGATATCAGCAATCTCTATCTGTTGGTGGCTGCGTACGCAATCCAGCGCGGCTGCC 5106
DB 4891 AAGCACATTGATAATCTTATCTGTTGGCGCAGCACCCTCTGGCGCGGCATCC 4950
QY 5107 GCGGTATCGTTGCGGCAAGCGCATGCGCAGCGCTGATGCTGGAGATGCGCCGCAATGA 5166
DB 4951 GCGCTAATCGGTCGCGCAAGCGCATGCGCAGCTTAATGCTGGAGGACCTGATTTGACGA 5010
QY 5167 ATCGACAGCTTTTACTTGAAGCAAGTAAACGCAACCATGCGGTGGCTCGAAGATTTG 5226
DB 5011 AT---ACGTCATATCTGAATCATGCGCTGAAACCATGCGGTTGGCTCGAAAGCTTTG 5067
QY 5227 CCACCGCCCGCAAGCTTTGATGACCGACGCGCGCAGCAGCTGATGCTGTATGCGT 5286
DB 5068 CGACTGCTCGACGCTTTTCGACGCGCAAAACCGCTCGCAGCGTGTGATGCTTTACGAT 5127
QY 5287 GGTGTGCTCACTCGGATGATGATGATGATGGGCAAAACGCTGGCGAAGCGCGCACGACG 5346
DB 5128 GGTGCGGCACTCGCAGCAGCTCAATTGACGATCAAAACACTGGGCTTTTCATGCGCCAGC 5187
QY 5347 ATGCGGTGCAAGACGCGCAGCAGCTATGACGATCTGCAAAATGAAACCGCGCGCT 5406
DB 5188 CTTCTTCGAGATGCTGAGCAGCGCTCAGACGCTTGAATGAAACCGCTCAGGCT 5247
QY 5407 ACAGCGCGCGCACATGATGAACCGCGCTTTAGGCGCTTCAGGAAGTGGCGATCAATC 5466
DB 5248 ACGCGGTTTCGAAATGACAGCGCGCTTTTGGCGGCTTCAGGAGGTTCGCGATGGCG 5307
QY 5467 ACCAGTCCCGCAACAACTGGGCTTTGATCATCTGGAAGGCTTCGATGATGATGACGCA 5526
DB 5308 ATGATATCGCTCCCGCTTACGGGTTTCGACCATCTGGAAGGTTTTCGATGATGTCGCG 5367
QY 5527 ACGAATTCACGAGCTTCGATGACACGCTGCTGCTATCATCGTCGCGCGGCTGG 5586
DB 5368 AAACGCGCTACCTGACACTGAGAGATACGCTCGGTTTATGCTATCACGTCGCGGTTG 5427
QY 5587 TCGGTTTGTATGTCGCGCGCTAAATGGCGGTGCGGACGAAGCGGTTCGATCAACGCT 5646
DB 5428 TGGGCTGATGATGGCGCAATTAATGGGCTTCGCGATTAACGCCACGCTCGAGCT 5487
QY 5647 GCGATTTAGGACTGGCGTTCCAGCTCACTAAATTTGCGCGCGACATTTGTAGAAGATGCGG 5706

DB 5488 GCGATCTCGGCTGCTTTCAGTTGACCAATTCGCGGTGATATTGTGACGATGCTC 5547
QY 5707 AAAATGGTCTGCTGCTATTCGCCGAATCTCGCTCGATCAGCGGGATTAACGCGCGATA 5766
DB 5548 AGTGGGCGCTGTTATCTGCTGAAGCTGCTGGAAGAGAGGACTGACGAAAGCGA 5607
QY 5767 CGTGACTGACCGCAACATCGTGAGCGCTCGCTCACTGGCAGCGGCTTTAGTGGCG 5826
DB 5608 ATTATGCTGCGCCAGAAACCGGAGGCTTTAAGCCGTATCGCCGGGAGCTTGTACGGG 5667
QY 5827 AGGCGAAACCTTATTCACCTCGCGCGATCCGCTTACCGGTTTACCGCTGCGCTCG 5886
DB 5668 AAGCGAAACCTTATTCATCATCAATGGCGGCTGCGCAATTAACCTTACGCTCG 5727
QY 5887 CGTGGCCATCGCTACCGCTCGCGCTTTATCGCAAAATTTGGCGTCAAAAGTTTCAGCAG 5946
DB 5728 CTTGGCCATCGCGACGAGCAGGTTGTACCGTAAATTTGGCGTGAAGTTGAACAG 5787
QY 5947 CCGTGTGCACGCTCGGATTTCACGGAGCGCACAGTAAAGTTGAAAACTGGGCTGC 6006
DB 5788 CCGTAAAGCAGCGCTGGGATCATCGCCAGTCCACGCTCCACCGCGAAATTTAACGCTTT 5847
QY 6007 TGCTGAAAGGCGAGGTTTGGGATCACTTCGCGTGTCTCGTCTGAAACCGCTGCGG 6066
DB 5848 TGCTGACGGCATCCGCTCAGGAGTTACTTCCGATGAAGACGATTCACCCCGCTCG 5907
QY 6067 CTGTGCTGTGGCAGCGCTCTCTGTTTATCTCGTGGTACGCTGGCGCAGCGTGGCTTG 6126
DB 5908 CTCTCTCTGGCAGCGCC-----GATCTAGCGCATGCTTCTCTCAGCGTGGCTG 5961
QY 6127 CAGCTTATTCAGCGGTGGCGGTAGAGGAAACCAACGACGACGAGCTTTCAGCGCGCG 6186
DB 5962 AAGTTTAGATAACCGTGGCGGTACAGAAACCAAGGACACGCGAGCTCTTTTCCCT 6021
QY 6187 CACCGCATGATGCTGCGGTGGCGCATCTAAGCGCTTAAGATAGCTTTTGGCGGGAT 6246
DB 6022 TACAGCATGATGCTAAGCGGTGGCGCATCTAAGCGTTTCAGGTAGCTTTTGGCGGTAT 6081
QY 6247 ATAGCGGAACCGCGAGCTTGTATGACACGAGCGCATCTGTCACCATGAAGTAGAGCGCG 6306
DB 6082 GTAGCGGAACCGCGAGCGCTGTTGATCCAGTCCGTCGTGGACCATATAAATACAGTAAAC 6141
QY 6307 GTACGCTGCTATTCGGGACCAATCCACTGACGCGGCGCACATGCTTGCACACGACATA 6366
DB 6142 ATAGCGGCTCATGCTGCAACCAATCCACTGGAGCGCGCAGATTCCTGTACTGCGCAAGTA 6201
QY 6367 AATCAGCAATCGCAGTACCGCAACACCAACCGCATAAAGATCGTTGAGCTCAAACTT 6426
DB 6202 AATCAGGCAATCGACACAATGGCGAATACCAACGCGATAGATCGTTAACTTCAATGC 6261
QY 6427 ACCGCTGTGCGGTTTCATGTCGACAGATGCGAGCGCCCATCCCAACCGTGCATGATGA 6486
DB 6262 GCTTTACGCGGTTTCATGATGTAAGATGCCAGCCCAACCCAGCGGTGCATGATGA 6321
QY 6487 TTTATGCAACGCGCGCTACGATTTCCATCACCAACCGGTTGCAACAGATAGCAC 6546
DB 6322 TTTATGTCGAGTGACCAACCACTTCCATGCGCACCGGTGACAAACACGATCAGGCG 6381
QY 6547 GTTCCATTAACAGAGCATTTGTTGCTTCATT 6576
DB 6382 ATTCCAAATCCAAACATAAATCTCTCCAGT 6411

RESULT 5

ERWCRTA

LOCUS

DEFINITION

ERWCRTA 6965 bp DNA linear BCT 11-APR-2001
Pantoea agglomerans CrtE (crtE), CrtX (crtX), dycopene cyclase
(crtY), phytoene dehydrogenase (crtI), and prephytoene
pyrophosphate synthase (crtB) genes, complete cds.

ACCESSION

M90698

VERSION

M90698.1

GI:148393

KEYWORDS

QY 876 CAACGAACTGAAACCGAGCGTCTGTTTCGCGCCACGCTGCAAAATGGCGGCGAATTGCCGC 935
DB 1480 GAAACACTTTAAACACGACACCTGTTTGGCGGTCTATAGCAGATGGCGCTCAATTTGTGGC 1539
QY 936 TGACGCTTCCACGCGAGTGGCGAAAGACTTACGCTTCTTCGCGCCAGGATTTGGCGCAGGC 995
DB 1540 GAATGCCCTCAGGGAACGAGCGATTAACCTACACCGTTTTCGCTTGATCTCGGCAAGC 1599
QY 996 GTTTCAACTGCTCGACGACTCGCGACCGTGGTGGAAACACACCGGTAAAGATGTGACCA 1055
DB 1600 GTTTCAACTGCTCGACGACTTAACCGATGGCATGCTGACACGGGTAAAGACATCAATCA 1659
QY 1056 GGATCAGGCAATCAACGCTGGTACAGATGCTCGGTGCTGACGCGGGAACGTCGCT 1115
DB 1660 GGATGAAGGCAATGACCGCTGGTCAATTTATAGGCTCCAGAGCAGTTGAAGAGCGTCT 1719
QY 1116 GCGCATCACTCGCGCAGCGCAGATGACACCTTCTCGCGCTGCACTCGCGGATCGC 1175
DB 1720 GCGGCACAGTCTGCACTCGCCAGTGAGCACTCTCAGCGCGGTGTCAAAACGGTCATTC 1779
QY 1176 CACTCGCCAAATATGACGCGCTGTTTAAATCAACAGCTPAGCGATTAATCAACTGACGCG 1235
DB 1780 TACTCAACAAATTTATCCAGCGCTGTTTGACAAAACCTCGCTGCGCTCAGTTAAGGATG 1839
QY 1236 GCTCAGCGGTGGCGCACTTTCGGGTGATCGCGCGCGCTCTACAGCACTTTCAGCG 1295
DB 1840 CTGCA-----TGAGCCACTTCGCGGATCGCTCGCGCTTTTACAGCCCATGCTCGTCA 1894
QY 1296 TTGCAGCGGTAGCAAAACGCTGTGGCGCGCGCATCGCATCACATTTATCAGCAA 1355
DB 1895 CTGGAACCTCGCTCAGAGTTAGTGGCTCGCGGTATCAGGTGACCTTTATTCAGNA 1954
QY 1356 GCGGATGCGCGCACTTTGCTTAGCGACGAACCGATCGATTTTGTTCGCGTGGCCAAACAG 1415
DB 1955 CAGGACACTAAACCACTCATTCAGCATCCCACTCGGATTTTCGTGTAGGAGAGAA 2014
QY 1416 AGCATCTGCGGTTCGCTGCGCGCGCTGTTGCTATCGCGTGGCTCGCGGCGCGCTG 1475
DB 2015 AGCCATCGCTGCGCTCAATTTACGCGCTGCTGCACTTAGCGGCCCATCTCTGGGGCG 2074
QY 1476 TCGCTGTTTCGCGTATCAACATCTCGCGTCTCGCACCGATATGCTGTCGCGCACTG 1535
DB 2075 TCAATGCTTCGCTCATCAATGAATGGCGCGCATCAGGAGATATGCTGTCCTGCGCTC 2134
QY 1536 CTGCGGTACTGAAGCAATGAACATCGATGCGGTGATCGCGACGAATAAGGAGCGG 1595
DB 2135 CCAGAGCGCTTAAACGATTTGCGATAGATGGTTCATCGTTGATCAGATGAGCGACGA 2194
QY 1596 GCGGATGCTGCTGAAGCGCTGCATCTGCGGTGTTGTTGCGTGGCTCGCGCTGCGG 1655
DB 2195 GCGCGCTGCTGCTGAAGCGCTGGGGCTGCGGCTTGTCTCGGTTGCTGTCACCTCCG 2254
QY 1656 GTCAATCGTGAAGCGGGATTCGCTTGGGTGATGCGCTTCGCTTTTGCACAGGATGAC 1715
DB 2255 CTTAATCGGAACCGGATTTGCCCTTGGCGGTGATGCGCTTTTGAATGCGACAGCGAT 2314
QY 1716 AAAGCGCTGAACGTTTTCAGGCGCAGCAGCATATCTATGATCGCATCATGCTGCTCAC 1775
DB 2315 GCGCAGCTGAGCGCTATGACGCCAGCGAAATTTATGATCGGTGATGCTGCGCAT 2374
QY 1776 GCGGAGTGATCTCAACACGCGGCGGTGTTTAAATTTGACGAGGCGGCGGATTAAT 1835
DB 2375 GACCGGCTATTTGCCCATCACAGTACAGGATGGGCTTAGCACCCCGGGATAAATCCAT 2434
QY 1836 CAGTGCCTGCGCGCTGCAAAATCAGCAGATGCTGCGCGCTTTGATTTTCCAGCT 1895
DB 2435 CAGTGTGTTTTCGCAATGGCAAAATCAGTACGCTTATTTCCAGAACTGATTTTCCCGT 2494
QY 1896 CAGCAACTCGCGCTGCTATCAGCGCTGGGGCCACTCCGCGCGCGCTTCTCTGCG 1955
DB 2495 AAAGCGCTACCGCTGTTTTCATGCTGTGGCGCTTGGCGCGCTTGGAGAAACGACAGCATCA 2554
QY 1956 CCGTCCATGCGCGCTGGCCAGCGCTGCTGCT---CAGCGGTGGTATATGCTGCTGCTGCGT 2012

DB 2555 ACCTTTGCGCGCTTATTTTTCGCGCTGCAAAAGCCCGAATTTTCGCTCGCTGGGC 2614
QY 2013 ACCTGCAAGGCAATCGCTTCGCGCTGTTTCTGCACTCTGGCGCAGCGCTGCCCGCAGCTG 2072
DB 2615 ACCTTGCAGGCTCATCGCTACGGTCTGTTTAAAGCCATCGTGAAGCCCTGTGAAGAGATT 2674
QY 2073 CCGCTATCGCTGCTGATCGCCATTTGCGGGATTAAGCGCGAAACAGACGATCAGCTG 2132
DB 2675 GATGCTCAGCTTCTGTTGGCGCACTGTGGCGCTGCTGCTTTCAGGCGAAGAGCTG 2734
QY 2133 GAGCTCGCTGCGCGCGCTGCTGACGATTTGCTGATCAGCGCGCAGCCCTACAGCAC 2192
DB 2735 GCGCAAGCCGTCAACGACAGTTTGTGGATTTTGGCGATCAGTCCGCGCGCTGTCGAG 2794
QY 2193 GCGCAGCTGTTTATCAGCTCATCGCGGTAAACAGCGCGCTGGAAAGCACTTGAATTCGCT 2252
DB 2795 GCGCAGCTTGCCATTAACCATGGCGGATGAATACGGTCTGGATGCTTAATTTATCGC 2854
QY 2253 AGCGCGATGCTGCGCTGCGGATTTGCTTTGATCAGCGCGGCTGGCGGCGGCAATTGAG 2312
DB 2855 AGCGCGCTTTAGCGCTTTCGCTTGGCATTTGATCAGCGCGCGCTGCGCTCACGTTATGTT 2914
QY 2313 TGGCATGACGTTGCTGCGCGCACTACGCTTTAGCCGTTTATCATCACTGAGCAGCAT 2372
DB 2915 TCTCATGGGATAGGAAACGCTGCTCCGTTTTTACACAGCCATGCGCTGCGCGCCAC 2974
QY 2373 CTGCAACAGCTGCTGACCGAGCATGTTTACGCGCTTACGAGTGTACGAGTTCAGCGCGAC 2432
DB 2975 ATGCGTGCATTTGTCACCAACGTCGACTACCGCAACGATGACGAAATCCAGACCGCC 3034
QY 2433 CTGACGCGCAGAGCGGTTTTCAGCGCTGCGCGACATCGTTCGAGCAGCGCTGTGCGCAG 2492
DB 3035 CTTCGTTTTCAGCGCGCGCAATGCGCTGCGATATCGTTGAGCAGCGCATGCG-CAC 3093
QY 2493 CAGCAAGCTGCTGCGCGAGCGACCTGATGCGCACGCAATACCATGTTGTTTGGTTCG 2552
DB 3094 CGCGCAGCGCTTTTGAAGGGAGCGGCTATGCGACCGCATTTGATCTGATTTTAGTGG 3153
QY 2553 GTGCTGACTGCGCAATGCTTGAATGCGCTGCGCTGCTGCTCAATTTGAGCGCACCACTGA 2612
DB 3154 GCGCGCGCTGCTTAATGCGCTCATTTGCCCTGCGCTTCAGCAGCAGCAACCTGCATGCG 3213
QY 2613 AATGCTGCTGCTGAGAGCGATGCGCATTCGCGCAGGCAATCATACCTGCTGCTTTCATC 2672
DB 3214 GTATTCGCTTATCGAAGCGCGCGCGGCGGTAAATCATACCTGCTGCTTTCAGC 3273
QY 2673 ACAGCGATCTCAGCGCGCAACAACTTCGCTGCTGCAACCGCTGATTAACCGTGGT 2732
DB 3274 ACCTGATTTAACCGAGAGCAACATCGCTGGGTAGCACCGCTGGTGTATATCACTGCG 3333
QY 2733 CAGGTTATCAGGTGCGTTTCTGCGCTGCGCGCAATCTGGACGGGATTAATTTGTTCCA 2792
DB 3334 CCGACTATCAGGTGCGCTTTCCAGCGCGCTGCTGTAAGTGAACAGCGCTATTTTCTG 3393
QY 2793 TCGCATCAGCGGATTTTCCCGCCATCTTTAGCGCGCGATGGGTGACGATCTGTGACAA 2852
DB 3394 TGACTTCGAGCGTTTTCGCGAGGTTTACAACAGAGTTTGGTCAGCACTTATGGAATTA 3453
QY 2853 ACACAGCGTACACAGGTAAACCAACGAGGTGACGCTGCGGATGCGCGTGAACCTG 2912
DB 3454 GCGCGCGGTGCGAGAGTTTCAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3513
QY 2913 CTGCGCAAGTGTGATTTGATGCTGCGCGCTGCGCGCGCATCTGACCGGATTAATTTGCTGCTGCT 2972
DB 3514 GTGCGCGCGGTGATTTGATGGCGAGGTTATACGCCAAATTCAGCGCTGACAGTGGAT 3573
QY 2973 ATCAGTGTCTTGAACAAGAGTGGCAGCTGCGCGCAGCGCGCTGACAGCGCA 3032
DB 3574 TCCAGCGCTTATCGACAGGAAATGGCGACTCAGTAAACCGCACGCTTATCTTACCCA 3633
QY 3033 TCCTGATGATGCGCGCTGCTGATCAGCAAGCGGTTTATCGTTTCTTCTACAGCTGCGC 3092

Db 3634 TTATATTGATGCCAGCGTCGATCAGCAAAATGGTTATCGCTTTGTTTACAGCCTGCCGC 3693
Qy 3093 TCAGCGCGGATCGGCTATTGATTGAAGATACCCATTACGTTAAACAGCCCGCGTGGCG 3152
Db 3694 TATCTGCGACAGAACTGTTAATCGAAGATACCCATATATCATATCGATAACGGGACACTGGAAC 3753
Qy 3153 AGAACACCGCTCGTCAGCACATCGCCGACTATGCCAATCAGCAAGGCTGGAGCTGAGTA 3212
Db 3754 CTGAACGGCGCGCAAAATAATTCGCGATTATGCTGCCAGCAGGATTGGCAGCCTTCAGA 3813
Qy 3213 CGCTCTGGTGAAGAGACGCGCATATTACGATTACCTGAGCGGCAACATCGATCGAT 3272
Db 3814 CTCTGCTTCGTGAGGAACAGGCGCATTCGCCAATAACCTTAACGGCGGATAGCAACCGGT 3873
Qy 3273 TCTGCAACAGCAGCGCGGCCAAGCGTCGAGCGGCTCGCGCGCGGCTGTTTCAATGCCA 3332
Db 3874 TTTGGCAGCAACAGC ---CTTTAGCTGCGAGTGGACTGGTGGCGGACTGTTTCAACCCA 3930
Qy 3333 CCACCGGTTACTCTTGGCGTCCGCTGGCGCTAGCGAGTGTGTAGCAGCGCTGTTGC 3392
Db 3931 CAACCGGTTATTCACTCCCGCTTGGCTGGCGGATCGCGCTGAGTGGCTTGATG 3990
Qy 3393 CCACCGATGCCCTCAGCTCAGCCAAACATATCGAAGCTTTGCCCGTCAGCAGTGGCGG 3452
Db 3991 TCTTTACATCGTCTCAATTATCATAGGCGATTACCCACTTTTGCACCAGCGCTGGCAGC 4050
Qy 3453 AACAGCGATTTTTCGCTCTGCTAAACCGCATGCTGTTTGGCGGTAAGCGCGAGCAGC 3512
Db 4051 AGCAAGGCTTTTTCGCGATGTTGATCGCATGCTGTTTGGCGGATTTAGCGGACCCGCGGATTAC 4110
Qy 3513 GCTGCGGCTGATGCAACGTTTTTACCAGCTCGATGCCGGGTTAATTAGCGCTTTTACG 3572
Db 4111 GCTGCGGCTGTTATGACGCGCTTCTATGTTTACCTGAAGATTAACTCCCGCTTTTATG 4170
Qy 3573 CGGCGCAACTGCGCTGCGGATTAACCGGATTTCTGCGGCAAGCGCGCGGTGCCCA 3632
Db 4171 CGGCGAAACTCAGCTGACCGGCTACGTTAATCTGAGCGGCAAGCGCGCTGTTCCGG 4230
Qy 3633 TCGGTGAAGCGCTGCGCGCTGTTGAATTCTGTCGAACAGGGGAAGAAATAAGAACG 3692
Db 4231 TATTAGCGCATTTGACGGAATTATGACGACTCATCGTTAAGACGCAACATGAAACC 4290
Qy 3693 CACTTATGATTTGCGGAGGCTTTTGGCGGCTGCGCTGCGGATTCGCTGCAAGCGGC 3752
Db 4291 AACTACGGTAAATTGCTGACGCTTTTGGTGGCTGCGCATTAGCAATTGCTGCGAGCGGC 4350
Qy 3753 GGGCATACCAACACTTACTCGACGCGGCAACACGGGCGGACGCGCTATGCTT 3812
Db 4351 GGGGATCCCTGCTTACTGCTTGAGCAACGCGACAAACCCGCTGGCGGCTTATGCTA 4410
Qy 3813 TGAGGACAGTGGCTTTTACCTTCGATGCCGACCCACCGGTGATCACCGATCCGAGCGCAT 3872
Db 4411 TGAAGATCAGGATTCACCTTTTGTATGACGGGCTACCGTTATTACCGATCCGAGCGCAT 4470
Qy 3873 CGAAGATGTTTACGCTGGCAGGAAATTCGCTCAGCGATTACGTCGAGCTGATCCCGGT 3932
Db 4471 TGAAGAACTCTTACCTTGGCGGGAACACAGTTAAAGATTACGTTGAATGCTCGCGT 4530
Qy 3933 AACGCCCTTATCGCTGCTGCGGAGATGGCAACAGCTTGATTAGCAATAATCA 3992
Db 4531 TCGCGGTTTATCGCTGCTGTTGGAGTTCAGGAAGGTTTTCAAATTACGATAACGATCA 4590
Qy 3993 GCGCTGCTGAGCAGCAGATCGCACGTTTCAATCCGCAAGATGTAGAAGGCTATCGTCA 4052
Db 4591 GGCACAGCTTGAGGCGCAGATTACGAGTTTATCCACGCGATTGTGAAGGCTATCGTCA 4650
Qy 4053 ATTTCTGCTTATCAGGTGAAGTATTTAGAGAGGTTATCTGAATTCGCGACCGGTGCC 4112
Db 4651 GTTCTGAGTATTCACGCTGCGGTGTTTAAAGAGGCTATCTGAAGCTTGGCACCGTGCC 4710
Qy 4113 GTTCTGAGGTCGCTGATCATCTCGCGTTCGCGCGGAGTGGGACGCTTGTGCAAGCATG 4172
Db 4711 TTTCTTGTGCTCAGGATATGCTTCGCGCGCGCCCCCAACTGGCGAAACTGCGAGGCATG 4770

Qy 4173 GCGAGCGCTTACAGCATGCTGGCGAAATTTATTAGGACGATCATCTGCGTCAGCGCTT 4232
Db 4771 GCGAACCGTTTACAGTAAAGTTGCGAGTCAATTAGGATGAACATCTTCTGTCAGCGCTT 4830
Qy 4233 TTCTTCCACTCATTTGCTGGTGGCGGTAATCTTTTTCGCAACGTCATCGATCTATACCTT 4292
Db 4831 TTTCTTTTCACTCGCTGTTAGTGGCGGCAATCCCTTTCGACCTCTCTCCATTTATACGCT 4890
Qy 4293 AATTATGCTGAGCGTGAATGGGCGTGTGTTTCCGCGCGGCGCAACGCGCGCT 4352
Db 4891 AATAATGCACTCGAAACGCAATGGGCGTGTGTTTCCACGCGGTGATTCGCGGCAAT 4950
Qy 4353 GGTGAGGCGCATGGCGGCTGTTTCGAGGACTGTGGCGCGGAGCTGTTACTGAAATGCCGA 4412
Db 4951 AGTGAAGGATGATAAAGCTGTTTTCAGGACTGGTGGCGAAGTGGTACTGNAATGCCAA 5010
Qy 4413 AGTGAAGCGCTGGAACACAGCGGCAATCGCAATTAGCGGCGTTCAGTTAGAGGCGGACG 4472
Db 5011 GGTGAGCCCATGGAACACAGCGGCGATACCAATTGAAGCGGTGCAATTTAGAGGACGAGC 5070
Qy 4473 ACGCTTGCATCGCGCGCTGTCCTCAATCCGAGCGTGGTGCATACCTACGACAACT 4532
Db 5071 CAGATTCCCGACCCGCGCTGTCGCTTCCAAATGCGGATGTGTTTCACTATCGGACCT 5130
Qy 4533 GCTTCGCCACCATCCGCTGGCAATGAACGCTGCGACATCGCTGAAGCGTAAAGCGCATGAG 4592
Db 5131 GTTAAGTCAGCATCTTGC CGCGGTGAAGCAGTCGAAAAAATGCAAACTAAGCGCATGAG 5190
Qy 4593 CAATCGCTGTTGTATCTCTAATTTTGGCTGAAATCAGCGGCAATGAACAGCTCGCGCACCA 4652
Db 5191 CAATCATTTGTTGTTCTCTAATTTTGGTCTGAATCAACCATCAGCATCAGCTTGCCACCA 5250
Qy 4653 CACGCTGTTTTCGCGCGCTATCGTGAAGTGCATCGATCAGATTTTCAACATCATGACGG 4712
Db 5251 TACGTTTGTTCGCGCGCTTCGCTACCGTGAAGTGCATGATCAATGAAATTTTCAATCATGACGG 5310
Qy 4713 GCTGCGACAGATTTTTCACCTTTTACCTGCGCGCGCTTGCAGCAGCGCATCGCTCGCTGCG 4772
Db 5311 ACTGCGACAGATTTCTCACTTTTACCTGCGCGCGCTTGTGTCAAGGATTCGTCACTGCG 5370
Qy 4773 ACCGCGCGCTGCGCGAGCTTTTATGTTTGAAGCGCGGTGCGGATCTCGGACCGCTGA 4832
Db 5371 ACCGGAAGTTTGCAGGCTTACTATGTTAGTCTCCGCTGCGACATTTAGGACAGCGAA 5430
Qy 4833 CATGACTGCGCAACAGGAAGCAGCGCTTCCGCGATCGAATTTTGTCTTATCTGAGCA 4892
Db 5431 CCTGACTGGAAGGTTGAAGGCGCGGACTCGCGGATCGTATTTTGTGATACCTTGAGCA 5490
Qy 4893 GCACCTACATGCGCGGATTTACGTCAGCAATTAGTGACACACAGAAATGTTTACGCGGTTTGA 4952
Db 5491 ACATTACATGCTGCTTTCGCTAGCCAGTTGTTTACGCGCGGATGTTTACGCGGTTTGA 5550
Qy 4953 TTTTCGCGACAGCTGCGATGCGGCTCGCGCTCGCGGTTTTCGCTGAGGACCGATTTTTCAC 5012
Db 5551 TTTTCGCGACAGCTCAATGCTTCAAGGCTCAGCTTTCGCTGAGCGGCTTCTCAC 5610
Qy 5013 GCAAGCGCTGTTTCGCGCGCATACCGCGATGCGGATATCAGCAATCTCTATCTGCT 5072
Db 5611 ACAGGCGCTTGGTTCCAGCGCAACCGCGATGAACCAATCAATAATCTTTACCTGCT 5670
Qy 5073 GGTGCGCGGTACGATCCAGCGCGGCGTCCCGCGGTGATCGGTTTCGCGCAAGCGCAC 5132
Db 5671 AGGTGACGACCCATCCCGCGCGGCTTCCCGCGGTGATCGGCTCGGCTAAGCGCAAC 5730
Qy 5133 GCGCAGGCTGATCTGAGGATCGCGCGGATGAATCGACAGCTTTTACTTGAAGCAAGTA 5192
Db 5731 TGCAGGTTTAAATGCTGGAGGA ---CCTGATTGAATAATCCGCTGCTTACTCAATCATCGG 5787
Qy 5193 ACGCAAACTATGCGGCTGCGGCTCGAAGAGTTTTCGCGCGCGCGCAAGCTGTTTTCATGCA 5252
Db 5788 GTCGAAACGATGCTGTTGGCTCGAAAGTTTTCGCGCTGCGCTCNAAGTTATTTGACGCA 5847

Qy	5253	CCGACGCGCGCAGCACGCTGATGCTGTATGCGTGTGTGCTCACTGCGATGATGTGATT	5311
Db	5848	AAAACCCGCGCGCAGTGTACTTATGCTCTACTCTCTGTGTGCCCATTTGTGACGACGTGATT	5907
Qy	5313	GATGGGCAACCGCTTGGCGGAAGCGGCACGCGACGATGCGCTCGAAGACGCGCAGGACGTT	5372
Db	5908	GAGCATCAAACCCCTGGGTGTTTCAACCGATCAGCCGTCATTACAAACGCTTGAGCAACGC	5967
Qy	5373	ATGCAGCATCTGCMAATTGAAACCCGCGCGGCTTACAGCGCGCGCACATGAGTGAACCG	5432
Db	5968	CTGATGCAGTTGGAGATGAAACCGCGCAGGCCCTATCCCGATGCGAGATGCGACRAACCG	6027
Qy	5433	CGGTTTAGGGCGTTTACAGGAAGTGGCGATCATTCACGAGTGGCGCAACAACTGGCGTTT	5492
Db	6028	CGGTTTCTGCGTTTTCAGGAAGTGGCGATGGCCCATGATATGCGCCCGCTTACGCGTTT	6087
Qy	5493	GATCATCTGGAAGGCTTCGCTATGGATGACGACGAAACAACTTACGCGAGCTTCGATGAC	5552
Db	6088	GATCATCTGGAAGGCTTTGCCATGGATGTGCGTGAGGCGGAGATACATTCACTAGTTGGACGAT	6147
Qy	5553	ACGCTGCGTTACTGCTATCACTCGCGGGCGTGTGCTGTTTGATGTGGCGCGCTGAATG	5612
Db	6148	ACATTACGGTATTGCTATCACTGTGCAAGCGTGTGCGCTTGATGTGGCGCAGATTATG	6207
Qy	5613	GGCGTGGCGACGAAGCGGTGCTCGATCAAGCTGTGCGATTTAGAACTGGCGTTCCAGCTC	5672
Db	6208	GGCGTACGCGATAACGCTACACTGGAACGCGCGCTGTGACCTTGGCGTGGCATTTCACTT	6267
Qy	5673	ACTTAACATTGGCGGACATGTAGAGATGCGGAAATGGTGCCTGCTATCTGCCGCAA	5732
Db	6268	ACCAACATCCCGCGATATTGTGGAAGATGCGCATGTGGGTGCTGTTTATCTGCGCGCG	6327
Qy	5733	TCCTGGCTCGATCAGGCGGATTAACGCGCGATACGCTGACTGCACCGCAACATCTGTGCA	5792
Db	6328	AGTTGGCTGGAAATGAAGNACTGAATAAGAAATTTATGGGCACCGGAAACCGCCAG	6387
Qy	5793	CGCGTCCGCTCACTGGCAGCGCGTTTAGTGGCGGAGCGGAAACCTTATTACTACCTGGCG	5852
Db	6388	CGCGTAGCGCGCATCGCGCGACGCTCTGCTGCAGGAAGCCGAAACCTTACTATTGTGCGCT	6447
Qy	5853	CGATCCGCTTACCGGGTTTACCGCTCGGCTCGCGGTGGCGCATCGCTACGGCTCGCGGC	5912
Db	6448	ACGCGGGGACTTGGCGGGGCTGCCCTTCGCTCAGCTTGGGCGCATTCGCGCAGGAAACG	6507
Qy	5913	GTTTATCGGAAATTTGGCGTCAAAAGTTTCAGCACGCGCGTGTGCACGCTCGGATTTACGG	5972
Db	6508	GTTTATCGTAATTTGGGTGAGGTTTGAACAGCGCAGAAACAGNAGCTCGGACCATCGG	6567
Qy	5973	CAGCGCACCAAGTAAAGTGA AAAACTCGCGCTGCTGTGTGAAGGGGAGGTTTGGCGATC	6032
Db	6568	CAGTCAACCAACAGACCCCGAAAAATTTATCGCTTCTGTGTGATGGCATCCGCTCAGGCCATT	6627
Qy	6033	ACTTCGCGTGTCTCGTCTCTGAACCGCGTCCGCGTGTGTGTGGCAGCGCTCTCTGTTGA	6092
Db	6628	ACTTCCCGGATGCGGCGCTCATTCGCGCTTCGCGCTGCCCATCTCTGGCAGCGCCGAA	6682
Qy	6093	TTTTACGCTCGTGCAGCTGGCGCAGCGTGGCTTCAGCTTATTTCAGCGGTGGCGCGTAGA	6152
Db	6683	TTTTAGCGGAATGTCTTCTTAGCGTTCGCTGAGGTTTTGTATAGCGCGGCGGCATAGA	6741
Qy	6153	GGAAACCAAAACGACAGCAGCTTTCACGCCCGCGCACCGCATGATGATCGGTGGCGCCA	6212
Db	6742	GMAAACCAAAAGAAACACAGCCTTCTTTGCTCTTTACGGCATGATGATCGGTGTGGGCCA	6801
Qy	6213	TGTATAGCGGTTAAGATAGCCTTTGGCGGGGATATAGCGGAAACCGGCACGCTTGATGCA	6272
Db	6802	TATACAAACCGCTTCAGGTAGCCTTTGGCGCGGAATATCCGGCAACGGGCAACGCTGTATGCA	6861
Qy	6273	CCAGGCCATCGTGCAACCATGAAGTATAGAGCGCGCGTACGCTCGTCTATTTCGGCACCAATCC	6332
Db	6862	CCAGCCGCTGTGAAACCATGAAATAGAGCAATTCATAGGCCGTCTATGCCCGCACCAATCC	6921
Qy	6333	ACTGCAGCGGCCACATGCCCTTGCAACCGGCATAAAATCAGCA 6374	

Db 6922 ACTGAAGCGCCACATTCCGTACTGCCACATAAATAGCA 6963

||||| ||||||||| || ||||||||| |||

RESULT 6
 ERWCRTS 12753 bp DNA linear BCT 11-APR-2001
 LOCUS Pantoea agglomerans pyrophosphate synthase (crtE), zeaxanthin
 DEFINITION glucosyl transferase (crtX), and lycopene cyclase (crtY) genes,
 complete cds; phytoene dehydrogenase (crtI) gene, partial cds;
 phytoene synthase (crtB) and beta-carotene hydroxylase (crtZ)
 genes, complete cds; and unknown genes.
 ACCESSION M87280 M99707
 VERSION M87280.1 GI:148404
 KEYWORDS
 SOURCE
 ORGANISM
 Pantoea agglomerans
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Pantoea.
 REFERENCE 1 (bases 1 to 12753)
 AUTHORS Hundle, B., Alberti, M., Nievelstein, V., Beyer, P., Kleinig, H.,
 Armstrong, G.A., Burke, D.H. and Hearst, J.E.
 TITLE Functional assignment of Erwinia herbicola Eho10 carotenoid genes
 expressed in Escherichia coli
 JOURNAL Mol. Gen. Genet. 245 (4), 406-416 (1994)
 MEDLINE 95107236
 PUBMED 7808389
 REFERENCE 2 (bases 1 to 12753)
 AUTHORS Armstrong, G.A., Hundle, B.S. and Hearst, J.E.
 TITLE Evolutionary conservation and structural similarities of carotenoid
 biosynthesis gene products from photosynthetic and
 nonphotosynthetic organisms
 JOURNAL Meth. Enzymol. 214, 297-311 (1993)
 MEDLINE 93225848
 PUBMED 8469144
 REFERENCE 3 (bases 1 to 12753)
 AUTHORS Hundle, B.S., O'Brien, D.A., Beyer, P., Kleinig, H. and Hearst, J.E.
 TITLE In vitro expression and activity of lycopene cyclase and
 beta-carotene hydroxylase from Erwinia herbicola
 JOURNAL FEBS Lett. 315 (3), 329-334 (1993)
 MEDLINE 93138098
 PUBMED 8423226
 REFERENCE 4 (bases 1 to 12753)
 AUTHORS Armstrong, G.A., Alberti, M. and Hearst, J.E.
 TITLE Conserved enzymes mediate the early reactions of carotenoid
 biosynthesis in nonphotosynthetic and photosynthetic prokaryotes
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9975-9979 (1990)
 MEDLINE 91088634
 PUBMED 2263648
 REFERENCE 5 (sites)
 AUTHORS Hundle, B.S., O'Brien, D.A., Alberti, M., Beyer, P. and Hearst, J.E.
 TITLE Functional expression of zeaxanthin glucosyltransferase from
 Erwinia herbicola and a proposed uridine diphosphate binding site
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (19), 9321-9325 (1992)
 MEDLINE 93028456
 PUBMED 1409639
 FEATURES
 Location/Qualifiers
 1..12753
 /organism="Pantoea agglomerans"
 /mol_type="genomic DNA"
 /strain="Eho10"
 /db_xref="taxon:549"
 /tissue_lib="ATCC 39368"
 complement(1..1588)
 /note="ORF2"
 /citation=[1]
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AA064974.1"
 /db_xref="GI:551854"
 /translation="MSELIHANAWIEWRFERQILRLPEWGNKSLNLRVATCAPETFSAL
 HALLTPESQAILRDAETITLRNGNISAVLNKQLAFYNQRLLEEMWQRST"

CDS

```
GIGASEKSDKYVSALKLDGREFKPLPGCKYQLTVRFPEARQCEKLYGMGOYQQPWLDL
KGCSELAQRNSQAVPVPQSSGLGYGLLWNNPAIGASFAXNKSOWTSRVTVQRNPLS
TAGDSMEITROYAKATCTPPAAPPEFIISGLWCKLRVYRTOEVLDAVEAREHNNPLS
VWVIDPFWPNQGTWCFDPVDPMPDPRAMVELESNGIRIMVSVMPVTEARSPLEPOMK
AKGLVTSDRGVQVNLDPNGNTTFPDATHPQARQVWTVTKNYPDYKIGLFWLDEAE
FEYRAYDSDVTHAGPVLVEGNRYPRDFPAQGFIDGLRAEGENEIVNLVRCWSWAGSOR
YGVLAWSGDVSHFSFRNQLAAGLNMGLAGIPWTTTIDGGFQGNTHDPAPHELLIR
WFQWAVFCPLRMHGYSRBPRI"
complement (1596..2414)
/note="ORF3"
/citation=[1]
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAA64975.1"
/db_xref="GI:148406"
/translation="MLTWNNLLMVQLRQHRWCLFGVAYIFSSVGNGLTQTLILG
QLLRWHAPSTUTLMLATLPGFISGLGAWMYEVTYILGSASYFMWMLAAHILGKA
AGSLAKRLTRFNKVFQIFGYCAVLGASIALFPAPKSVFVLVELTFIISTLYQATT
TLMWYMAADVADYGSQGRMDGIIFSTFLAVLKLGMALGAIWGTILGFSGYVANA
PEOTSLMHCIVALFTVTVGGLSLAAAFATLRWYKLDQTMQEBINLAKMQTIYKG"
3140..3451
/note="ORF4"
/citation=[1]
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAA64976.1"
/db_xref="GI:148407"
/translation="VALRSLSGTTGLHIPPLVFGNVGFWTVDEQSFSLDALLERGL
NAIDTADYSAWAPCNKGESETILGKWFAPAHGAREKITLFTKVGSLDKAQLRLMLRT
M"
3535..4458
/genes="crte"
3535..4458
/genes="crte"
/citation=[4]
/citation=[2]
/codon_start=1
/transl_table=1
/product="pyrophosphate synthase"
/protein_id="AAA64977.1"
/db_xref="GI:148408"
/translation="MVVSGSKAGVSPHREIEVMRQSIDDHLAGLLPETSDQIVSLAMR
LRWAPGRKIRPELMLAARDLYQGSMTLIDLACAVELHTASLMDMPCMDNAE
EGMGQPTTKRFGESVAILASGYLSKAFGLIAAGDLPGERRAQVANELSTAVGQAG
LVLGQFRDLNDALDRTPDAILSTNHLKTLGILFSAMLOIVAIASASPSTRETLHAPA
LDPGAQQLDLRDDHDTGKDRNKDGKSTLVNRLGADAAARQKLRHIDSADKHLT
FACPGGAIQRQFMHLWFHGLHADWSFVWKIA"
4521..5564
/note="gene not found in Erwinia uredovora crt gene
cluster: ORF6"
/citation=[1]
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAA64978.1"
/db_xref="GI:148409"
/translation="MKDELVORKNDHLDIVLDPRAVTOASAGFERWFTHCALPEL
NFSDTLETFFLNROQLAPLLISSMTGGVRSRHNRHLAEAAQVLIKIAMGVGSORVA
IESDAGLDKTLROLAPDVPILLAGAAQLTGKRGIDYARRAVEMISADALIVHLNP
LQAGLPQGGDRWRGLAAETLVRELFPVLVVEVGAGISRTVAGQLIDAGVTVIDY
AGAGTSWAAVEGERAAEQQRSVANFADWGIPTAEALVDIAEAWPQMPLIASGKIK
NGVDAAKALRLGACMVGQAAAVLSGAGVSTEKVIDHFNVIIQLRVACFCCTGSRSLSD
LKQADIRVYRDTP"
5561..6802
/genes="crtx"
5561..6802
/genes="crtx"
/citation=[3]
/citation=[1]
/codon_start=1
/transl_table=1
```

```
/product="zeaxanthin glucosyl transferase"
/protein_id="AAA64979.1"
/db_xref="GI:148410"
/translation="MSHPAIVAPPLYSHAVAHALALEMAQORHVRVTLTGNVASLAE
QETSERVAFYPLPASVQQAQRNVQOQSGNGLRLIAAMSSLTDLVLCQOLPAIILQRLAVD
ALIVDEMPAGSLVAEALGLPFIACALPVNREPGLPLPVMPHYAEDKALRRFQV
SHYIDALMYPHGGQITILRHAQRFGLPERRRDLDECLPLAQISQVFPALDFPRRALPNC
FERYGALRYOPPOVERSPPRSTPRIFASLGTLOGRHLRFQKIRACASVGAETVIAH
CDGLTPAQAQSLYAGCTEVVDFVQPRYVAEANLVIHTGGLNTVLDAALAAATPVYLA
PLSFDOPAAARLVYNGLGRVSRFARQQTLLADEIAQLLGDETLHQLRLATARQQLNDA
GGTPRAATLLEQAIGSESUS"
6799..7959
/genes="crty"
6799..7959
/genes="crty"
/citation=[3]
/citation=[2]
/citation=[1]
/codon_start=1
/transl_table=1
/product="lycopenase cyclase"
/protein_id="AAA64980.1"
/db_xref="GI:148411"
/translation="MRDLILVGGGLANGLIARLQRYPOLNLLILIEAGEQPGGNHTW
SFHEDDLTPGQHAWLAPLVAHWPGYEVQFPDLRRRLARGYISITSERFADALHQAUG
ENIWLNGSVSEVLNSVRLANGEALLAGAVIDGRGVTAASSAMQTCYQLFLGQWRLTQ
PHGLTVPLMDATVAQQGVYFVYTLPLSADTLLIEDTRYANVPQRDDNALRQTVTDY
AHSGWOLAQLERETGCLPTTLAGIDQALWADAPGVPSGMRAGLPHPTTGYSLPLA
VALADATDSPRLGSLVPLQYLTROPAREHWERQGFERLLNMLFLAGREENRWVRQOR
FYGLPEPTVERFYAGRLSLEFKARILITGKPPVPLGEAWRAALNHFPPDRDKG"
7956..99414
/genes="crti"
7956..99414
/genes="crti"
/note="putative; same bases, 11..1489, Genbank M38423"
/citation=[4]
/citation=[2]
/codon_start=1
/transl_table=1
/product="phytoene dehydrogenase"
/protein_id="AAA64981.1"
/db_xref="GI:551855"
/translation="MKKTVVIGAGFGGLALAIRLOAQAGIPTVILLEQORDKPGRAYVWH
DOGTFDAGPTVITDPTALEALFTLAGRMEDYVRLLPVKPFYRLCWESGKITLDYAND
SABEAGITQPNRPDVEGRFPLAYSQAVFOEYRLGSLVFLFRDMLRAGPOLKL
QAWQSVQSVSRFTEDEHLRQAFSHLLVNGNFTTSSYITLTHALERGVWPFEG
GTGALVGMVKLFTDLGGEIELNARVEELVADNRKVSQVRLADGRIFTDVAASNADY
VNTYKLLGHPGVGKRAAALERKSMNSLFLVYFGLNQPHSLAHHTICRGPYREL
IDEIFTSALADDSLYLHSPCVTDPSLAPGCSFYVLAAPVPHLGNAPLDWAQEGPK
LRDRIFDYLEERYMPLGRSQVLTQRIFFPADFDHDLDAHLGSAFSIEPLLTQSAWFRP
HNRSDDIANLYLVAGTHPGAGIPGVVASAKATASL"
9431..10360
/genes="crtb"
9431..10360
/genes="crtb"
/note="same bases, 1486..2415, Genbank M38423 but now
function has been reassigned"
/citation=[4]
/citation=[2]
/codon_start=1
```

```
Query Match 30.5%; Score 2135.2; DB 1; Length 12753;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 3263; Conservative 0; Mismatches 1765; Indels 39; Gaps 4;

QY 1479 CTGTTTCGGTGATCGACGATCTCGCGTCCTGCGACCGATATGCTGCGCGGAAGTGCCT 1538
DB 5780 CTGTCGGCGGTGATTGCGGCATGTGTCCTCCCTCCGATGTCCTGCCAGAGTGTCCC 5839
QY 1539 GCGGTACTGAAAGCATTTGAACATCGATGCGGTGATCGCGCGAGAAATGGAAGCGCGGC 1598
DB 5840 GCTATTCTACAGCGGCTGCGGTGGACGCGTGTGATTGTCGATGAGATGAGTGGAGCGGCCGGA 5899
```


Db 8027 GATCCCAACCGTACTGTGGAGCAGCGGCAAGCCCGCGGTGGGCTACGTCTGGCA 8086
Qy 3816 GGACAGTGGCTTTACCTTCGATGCGGACCCACGGTGTACCCGATCCAGCGGCATCGA 3875
Db 8087 TGACCAGGGCTTTACCTTTGACGCGGGCCGACGGTGTATCCCGATCCATCCGCGCTGA 8146
Qy 3876 AGAGTTGTTCAGCTGGCAGGAAATCGCTCAGCGATTACGTTCGAGCTGATGCGGGTAAC 3935
Db 8147 GCGCTGTTCACCTTGGCGGCGAGCGCATGGAGATTACGTACAGGTGTCTGCGGTAAA 8206
Qy 3936 GCCCTTCTATCGCTGTCTGGAGATGGCAACAGCTTGATACGACAAATATACGCC 3995
Db 8207 ACCCTTCTACGACTCTGCTGGAGTCCGGAGAGACCCCTGACTATGCTAAGCAGACGCG 8266
Qy 3996 GCTGCTGAGCAGATGCGCAGCTTCAATCCCGCAAGATGATAGAGGCTATCGTCAAT 4055
Db 8267 CGAGCTTGAGGCGCAGATTACCGAGTTCAACCCCGCGACGTCGAGGGCTACCGCGCTT 8326
Qy 4056 TCTTGCCTATTACGCTGAAGTATTTAGAGAGGTTATCTGAAACTTCGGCAAGGTGCCGTT 4115
Db 8327 TCTGCTTACTCCAGGCGGTATTCAGAGGGATATTTGCGCTCGGACGGCTGCCGTT 8386
Qy 4116 TCTCAGGTGCTGATCATGCTGCGGTGCGCGCGAGTTGGAGCTGTGCAAGCANTGCG 4175
Db 8387 CCTCTCTTTTCGACATGCTGCGCGCGGCGCGAGCTGTCTTAAGCTCCAGCGCTGGCA 8446
Qy 4176 CAGCGCTACAGCATGGTGGCAATTTATTCAGGACGATCATCTGCGTCAAGGCGTTTC 4235
Db 8447 GAGCGCTACCAAGTCCGCTTTCGCGCTTATTTAGGATGAGCATCTGCGGAGGCCCTTCTC 8506
Qy 4236 CTTCCACTCATTTGCTGGTGGCGGTAAATCTTTTGAACGTCATCGATCTATACCTTAAT 4295
Db 8507 GTTCCACTCCCTGCTGTAGGCGCAACCCCTTACACCCTCGTCCATCTACACCCTGAT 8566
Qy 4296 TCATGCGTGBAGCTGTAATGGGCGGTGTGGTTTCGCGCGCGCGCAGCGCGCGCTGGT 4355
Db 8567 CCACGCGCTTGAGCGGAGTGGGGGCTGTGGTTCCCTGAGGGCGCACCGGGCGCGCTGGT 8626
Qy 4356 GCAGGCGATGGCGACTTTCGAGACTTGGCGCGGAGCTGTACTGAAATGCCGAAGT 4415
Db 8627 GAAAGGATGTGTGAGCTGTTTACCGATCTGGGCGGGAGATCGAACTCAAGCCCGGCT 8686
Qy 4416 GAGCAGCTGGAACACGCGCAATCGCAATTAGCGCGCTTCAGTTAGAGGGCGGACGACG 4475
Db 8687 CGAAGAGCTGTGTGGTGGCGATAA CCGCGTAAGCAGAGTCCGCGTGGCGGATGTGCGAT 8746
Qy 4476 CTTTCGATGCGCGCTGTGGCTCCATTCGCGAGTGTGTACCTACGACAACTGCT 4535
Db 8747 CTTTGACACCGACGCGGTAGCTCGAAACGCTGACGTGTGTGAACCTATAAAGCTGCT 8806
Qy 4536 TCGCCACCATCGCTGGCAATGAACGTCGACATCGCTGAAGCGTAAGCGCATGAGCAA 4595
Db 8807 CGGCCACCATCGGTGGGCGAGAGCGGGCGGACGCGCTGGAGCGCAGAGCATGAGCAA 8866
Qy 4596 CTGCTGTGTTGTACTCTATTTTGGCTGAAATCAGCGCGATGAACAGCTCGCGCACACAC 4655
Db 8867 CTGCTGTGTTGCTCTACTTCGGCTGAACACGAGCTCATTTCCAGCTGGCGCACCATAC 8926
Qy 4656 CGTCTGTTTTGGCGCGTTATGCTGAGTTGATCGATGAGATTTTCAACAGCAGCGAGCT 4715
Db 8927 CATCTGTTTGTGTCCTCGCTACCGGAGCTGTGACGAGATCTTTACCGGACGCGCT 8986
Qy 4716 GGCAGACGATTTTCACTTTACCTGACGCGCTTCAGCAGCGATTCCTGCTGCGACC 4775
Db 8987 GCGCGATGACTTCTGCTCTACCTGACCTCGCCCTCGCTGACCGATCCCTCGCTCGCGC 9046
Qy 4776 GCCCGCTGGGAGCTTTATGTTGTAGGCGGCTGGCGCATCTCGGACCGCTGACAT 4835
Db 9047 TCCCGCTGCGCAGCTTCTACGTGTGCGCGCGGTGCGCATCTTTGGCAAACGCGCGCT 9106
Qy 4836 CGACTGCAACAGGAAGGACCGGCTTGGCGCATCGAAATTTTGTCTTATCTGAGCAGCA 4895
Db 9107 GGACTGGCGCAGGAGGGCCGAAGCTGCGGACCGCATCTTGACTACTTTGAGAGCG 9166

Qy 4896 CTATATGCCGGATTACGTTCAGCAATTTAGTGACACAGAAATGTTTACCGCGTTGATTT 4955
Db 9167 CTATATGCCGGCTCGGTAGCAGCTGTGTGACCCAGCGGATCTTTACCCCGCAGACTT 9226
Qy 4956 TCGGCAGCAGCTGTATGCCCATCAGCGCTCGGGTGTTCGCTGGAGCGGATTTTGACGCA 5015
Db 9227 CCACGACAGCTGTGATCGCATCTGGGATCGGCTTCTCCATCGAGCGGCTGTGACCCA 9286
Qy 5016 AAGGCGCTGTTCCGCGCATTAACCGCATGCCGATATCAGCAATCTCTATCTGTGGG 5075
Db 9287 AAGCGCTGTGTTCCGCGCCACAAACCGCAGCAGCAATTTGCCAACTCTTACCTGTGGG 9346
Qy 5076 TGCCTGTACCATCCAGCGCGCGGTGCCCGGCGTGTTCGCTTCGSCCAAGGCCACCGC 5135
Db 9347 CGCAGGTACTCACCTTGGGCGGGCATTTCTGGGTAGTGGCTCGCGGAAAGCCACCGC 9406
Qy 5136 CAGCTGATGTGAGAGTCGCGCGCGAATGAATCGACAGCTTTACTTTGACGAAGTAACG 5195
Db 9407 CAGCTTGTATGAGGATC---TGCAATGAGCAACCGCGCTGTGTGACCGCACG 9463
Qy 5196 CAACACATGCGGTGGCTCGAAGATTTCCGACCGCGCCCAAGCTGTGTGATGCAACG 5255
Db 9464 CAGCCATGGCCAAAGGCTCGAAGATTTTGGCAACCGCTGGAGAGCTTTTCGACCGCGC 9523
Qy 5256 ACGCGCGCAGCAGCTGTATGCTGTATGCTGTGTGTCTCACTGCGATGATGTGAT 5315
Db 9524 ACCGCGTAGGTGTGTATGCTCTACACCTGTGTGCGCCACTGCGATGACGTCAATTGAC 9583
Qy 5316 GGGCAACGCTGGGCGAAGGGCGCACGACGATGCCCTCGAAGACGCGCAGGACGTAAG 5375
Db 9584 GACCAGACCCACGCGCTTTCGCGAGCGAGCGCGCGGAGGAGGACCAACCCAGCGCTG 9643
Qy 5376 CAGCATCTGCAAAATTGAAACCCCGCGCGCTTACAGCGCGCGCAGCATGAGTAACCGCG 5435
Db 9644 GCCCGCTGCGACGCTGACCCCTGGCGCGTGTGAAAGGGCCGAGATGACAGATCCGCGC 9703
Qy 5436 TTTAGGCGTTTACGAAAGTGGCGATCATTCACAGCTGCGGCAACAACTGGCGTTGAT 5495
Db 9704 TTCGCTGCTTTTACGAGGTGGCGCTGACCCACGCTATTACGCGCCCGCATGGCGCTCGAT 9763
Qy 5496 CATCTGGAAGGCTTCGCTATGAGTGCAGCAACAAACATTAACGCGAGCTTCGATGACACG 5555
Db 9764 CACCTCGACGCGCTTTCGATGAGCTGTGCTCAGACCCGCTATGTACCTTTGAGGATACG 9823
Qy 5556 CTGCGTTACTGCTATCAGCTCGCGGCTGTGCTGTGATGATGGCGCGGTAAATGGGC 5615
Db 9824 CTGCGTACTGCTATCAGCTGCGCGGTGTGCTGTGATGATGGCCAGGCTGATGGGC 9883
Qy 5616 GTGCGCAGCAGCGGTGCTCGATCACGCTGCGATTTAGGACTGGCGTTCAGCTCACT 5675
Db 9884 GTGCGGATGAGCGGTGCTGATCGCGCTTGGGCTTGGGCTTCCAGCTGACG 9943
Qy 5676 AACATTGCGCGCATTTGTAGAAGTGCAGAAATGCTGCTGTCTATCTCCGCAATCC 5735
Db 9944 AATATGCGCGCGATATTTATGACGATGCGCTATTGACCGCTGTCTATCTCCCGCGAG 10003
Qy 5736 TGGCTCGATCAGCGGGGATTACGCGCGGATACGCTGATGACCGCGCAACATGTCGACG 5795
Db 10004 TGGCTGAGGATCGCGGCTGACCCCGGAGAACTATGCGCGCGGAGAAATCGGCGCGC 10063
Qy 5796 CTCGCTCACTGGCAGCGGTTTAGTGGCGGAGCGGAACCTTATATCACTCGCGCGGA 5855
Db 10064 TGGCGCGGGTGGGAGCGGCTTATGATCGCGAGCGGTACTACTATCTCTCCAG 10123
Qy 5856 TCCGTTTTTACCGGGTTTACCGCTGCGCTCGGCTGGGCCATCGCTACCGCTCGCGCGCT 5915
Db 10124 GCCGGCTACAGATCTCGCGCGCTGCGCTGGGCGATCGCCACCGCGCGCAGCGCTC 10183
Qy 5916 TATCGGAAATTTGGCGTCAAAAGTTGACGACCGCGGTGTGACGCTGGGATTCAGCGGAG 5975
Db 10184 TACCGGAGATCGGTATTAAAGTAAAGCGCGGAGGACGCGCTGGGATCGCGCGCAG 10243

```
QY 5976 CGCACCAGTAAGCTGAAACCTGGCGTCTGCTGTAAGAGGGCAGGTTTGGCGATCACT 6035
Db 10244 CACACAGCAAAAGGTGAAAAATTCCTATGATGGCGGACCCGGGCGAGGTTATCGG 10303
QY 6036 TCGCGTGTCTGCTGCTGAAACCGCGTCCGGCTGCTGCTGCGCAGCGTCTCGTTGATTT 6095
Db 10304 GCGAGACGACGAGGTCAGCGCGGTCCGGCTGCTGCGCAGGTCCTTGGCAGGTCCTTGGCGG 10363
QY 6096 TAGCTCCGTGACGCTGGCGCAGCGTGGTGTGAGCTTATTCAGCGGTCGCGGTAGAGGA 6155
Db 10364 GCGCGGCTACGCTTCACGAGGATCGCTGAGTGGCGAGCTTGGCGGCGTAAATAA 10423
QY 6156 RACCAAAACACAGCAGCTTCACGCCCGCGCAGCGATGATGATGATGATGATGATGATGAT 6215
Db 10424 AACCAGAGAGAGCAGCGCCCTCCCGCGCGCAGCGGTCGAGCGGTCGAGCGT 10483
QY 6216 ATAAGCGCTTAAGATAGCTTTGCGCGGATATAGCGGAACCGCGAGCGTTCATGACCA 6275
Db 10484 AGAGCGCTTCAGGTAGCCCCGCGGGATCCAGTGGAGAGGCGCAGCGTATGACCA 10543
QY 6276 GGCATCTGTCACCATGAGTAGAGCGCGCGTACGTGTCATTCGCGCACCAATCCACT 6335
Db 10544 GACCGTCGTGACCAAGAGTAGAGCAGCGCATAGACCGTCATCGCGAGCCAAATCCACT 10603
QY 6336 GCAGCGGCCACATGCTTGCAACCGACATATAGCAGCAATCGCAGTACCGCAACA 6395
Db 10604 GCAGGGGCAACCGCGCGCGTCCCGCAATCAGCGCGATAGCCACCCCGGCAACA 10663
QY 6396 CCACCGCATAAAGATGTTGAGCTCAAACTTACCGCTGCGCGTTCATGTCGCGACAGAT 6455
Db 10664 CCACCGCAAGAGATCGTTAGCTCAATACCGCTTGGCGGGGTATGTTGATCAT 10723
QY 6456 GCCAGCCCATCCCAACCGTCATGATGATTTATGCGACAGCGCGCTACGATTTCCA 6515
Db 10724 GCCAGCGCATCCCAACCGTCATGATGATTTATGCGAGTGGGTAACCGCGGATGCCCTCCA 10783
QY 6516 TCACCACACAGGTTGCCACACAGATAA 6542
Db 10784 TCGCAATACGCTCAAGATGACGATTA 10810
```

```
RESULT 7
LOCUS CQ793208 4624 bp DNA linear PAT 29-MAR-2004
DEFINITION Sequence 24 from Patent WO2004018694.
ACCESSION CQ793208
VERSION CQ793208.1 GI:45824305
KEYWORDS
SOURCE Pantoea ananatis
ORGANISM Pantoea ananatis
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Pantoea.
REFERENCE
1 Sauer, M., Flachmann, R., Klebsattel, M. and Schopfer, C.R.
AUTHORS Method for producing ketocarotenoids in genetically modified
TITLES organisms
JOURNAL Patent: WO 2004018694-A 24 04-MAR-2004;
Sungene GmbH & Co. KGAA (DE)
FEATURES
Location/Qualifiers
1..4624
/organism="Pantoea ananatis"
/mol_type="unassigned DNA"
/db_xref="taxon:553"
128..>1267
/feature="unassigned protein product"
/codon_start=1
/transl_table=11
/db_xref="GI:45824306"
/transl_table=11
/translation="MPPHYDLILVAGLANGLIALRLOQOQDMRILLIDAAPOAGGN
HTWSPHDDLTQSHRWAPLVVHWPDPYVFRPFRRLKNSGYFCITSORFAEVLOR
QFPHLMDTVAEYNASVRLKQGVIGARAVIDRGVANSALSQVQAFIGEW
LSHPHLSPIIMDATVDQNGRFRFVYSLUPSLRLLIEDTHYIDNIDLDPECAKQNI
```

```
CDS
1288..2766
/feature="unassigned protein product"
/codon_start=1
/transl_table=11
/protein_id="CAG24422.1"
/db_xref="GI:45824307"
/translation="MKPTTVICAGFGGLALAIRQAAGIPVLLLRQDKPGRGAYVYE
DQGTFOAGPTVIDPSAIEBLFAGKOLKEVYVLLPVTVPYRLCWSMGKVFNDKL
QTRLEAOIQOENPRDVEGYRFLDYSAVFEKGYLKLCTVPFLSFRDLRAAPOLAKL
QANRSVSKVASYIEDHLRQAFSPHLLVGNPPATSIYTLIHALREWGVWPRG
GTGALVQGMKLFQDLGGVVLNARVSHMETKIEAVHLEDGRRFLITQAVSNADY
VHTYDRLLSQHPAAVKQSNLOTRKMSNLFVFLGNHHDQLAHTHTVCFPRYREL
IDEIFNHDGLAEFSLYLHAPCVLTHDSSLAPEGCGSYVYLAVPVHLGTANLWDVTEGPK
LRDIFAYLEQHYMPGLRSOLVTRMPTFPDQNLNAYHGSFVSEVPVLITQSAWFRP
HNRDKTITNLVLVGAGTHPGAGIPGVIGSAKATAGLMLEDLI"
2802..3692
/feature="unassigned protein product"
/codon_start=1
/transl_table=11
/protein_id="CAG24423.1"
/db_xref="GI:45824308"
/translation="MAVGSKSPATASKLFDKTRRSVLMYAWCHRCDDVDIDDTLGF
QARPALQTPERLQDLEMTKQYVAGSQMHEPAPAFQEVAMAHDIAPAYAFDHLG
FAMVREARQYQSLDDTLRYCYHVAGVVGWMAQIMGVNDNATLDRACDLGAFQJTN
ARDIVDAHAGRCYLPASKEHGLKENYAAENRQALSRIARLVEAEPEYLSAT
AGLGLPLRSAAJATAKOVYKIGVKVEQAQAWDQORQSTTTTPEKLTLLLAASQA
LTSRMAHPHPHMLMQRL"
3631..4158
idna
ORIGIN
```

```
Query Match 27.2%; Score 1900.6; DB 6; Length 4624;
Best Local Similarity 66.5%; Pred. No. 0;
Matches 2787; Conservative 0; Mismatches 1389; Indels 13; Gaps 4;
QY 2405 GCTACGATGTCAGGATTCAGGCGCAGCTGCGAGCGCGCGAGCGGTGCGAGCGTGC 2464
Db 12 GCAGCGCATGCGCAAAATCCAGACAGCCCTTCGTTTGGCAGGGGCGACCATGCGCGCTGC 71
QY 2465 GCACATCGTCAGCAGCGCGTGTGCGACGACGACGCTGCTGCGGAGGCGACCTGATG 2524
Db 72 CGATATCATTCAGCAGGTTATGTG-CACCGTCAGCCTGTCTTAAAGTGGAGCGGCTATG 130
QY 2525 CGCAGCAATACGATGTCATTTTGGTCGCTGCTGACCTGGCGAATGGCTTGAATGCGCTG 2584
Db 131 CAACCGCATTCATTCGATTCCTGCGGGGCTGACTCGCGAATGGCTTATCGCCCTG 190
QY 2585 CGTCTGCGTCAATTGCGACCACTGAAATGCTGTTGTCGAGAGCGGATCGGATCCG 2644
Db 191 CGTCTTCAGCAGCAGCAACCTGATGCGTATTTTGTCTTATCGACGCGCGACCCAGCGG 250
QY 2645 GCAGCATCATACCTGCTGCTTTCATCAGCGCATCTCAGCGCGGACCAACTTCGCTGG 2704
Db 251 GCGCGGAATCATACGCTGCTATTTCCACCATGATGATTTGATGAGAGCGCAACATCGTTGG 310
QY 2705 CTGCAACCGCTGATTCACGCTGCTTGGTCAGGTCAGGTCGCTTTTCTCGCTGCGCG 2764
Db 311 ATAGCTCCGCTGGTGTTCATCAGTGGCCCGCATTCAGTACGCTTTCCACACACGCCGT 370
QY 2765 CGCAATCTGGACGGGATATTTGTTCCATTCGATCAGCGGATTTTGGCCGCCCATCTTTAC 2824
Db 371 CGTAAGCTGAACAGCGGCTACTTTTGTATTTACTTCTCAGCGCTTTCGCTGAGGTTTTACAG 430
QY 2825 GCGCGATGGGTGACGATCTGTGGACAAACACAGCGGTACACAGTAAACCCAGCGAG 2884
Db 431 CGACAGTTTGGCCCCCATCTTGTGGATGATACCGCGGTGCGAGAGGTAATTCGGAATCT 490
QY 2885 GTGACGCTGGCGGATGGCGCGTGAACCTGCTCGCAAGTGGTGAATGATGCTGCGCGCTG 2944
Db 491 GTTCGTTGAAAAGGGTCAGGTTATCGTGCCCGCGGTGATTTGACGGGCGGGGTTAT 550
```


Db 2708 CTGGCGTCTATCGGCTCGGCAAAAGCGACGACGAGGTTTGATGCTGGAGATC---TGATTT 2764

Qy 5165 GAATCGACAGCCCTTTACTTTGACAAAGTAAACGAAACCATGGCGGTGGGCTCGAAGATTT 5224

Db 2765 GAATAATCGTCTTACTCAATCATGCGGTGAAACGATGGCAGTTGGCTCGAAGATTT 2824

Qy 5225 CGCACCGCGCAAGCTGTTTGATGCAACGAGCGCGCGCAGCAGCTGATGCTGTATGC 5284

Db 2825 TCGCAGAGCCTCAAAAGTTATTTGATGCAAAACCCCGCGCAGCTACTGATGCTCTACGC 2884

Qy 5285 GTGGTGTCTCATCTCGATGATGATGATGATGGGCAAAACGCTGGCGGAAGCGCGCAGCA 5344

Db 2885 CTGGTGGCGCATTTGTACGATGTTATTTGACGATCAGACGCTGGGCTTTTCAGCGCCGCGCA 2944

Qy 5345 GCATCGCGTGAAGACGCGCAGCAGTATGACGATCTGCAAAATGGAACCCCGCGCGC 5404

Db 2945 GCCTGCTTACAAACCGCGCAACACGCTCTGATGCAACTTGAGATGAAACCGCGCAGGC 3004

Qy 5405 CTACAGCGCGCACATGATGAACCGCGCTTTAGGCGGTTTCAGGAAGTGGCGATCAT 5464

Db 3005 CTATGAGATGCGAGATGCAAGAACCGCGCGTTTCGCGGCTTTTCAGGAAGTGGCTATGCG 3064

Qy 5465 TCACAGCTGCGCGCAACAACTGGCGTTTGATCATCTGGAAGCTTTCGCTATGATGATGCAG 5524

Db 3065 TCATGATATCGCGCGCTTACGCGTTTGATCATCTGGAAGCTTTCGCGATGATGATGACG 3124

Qy 5525 CAACGAACATTAACGAGCTTCGATGACACGCTGCGTTTCTGCTATCATCGTCGCGGCGT 5584

Db 3125 CGAAGCGCAATACAGCAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3184

Qy 5585 GTCGCTGTTGATGATGCGCGCTTAATGGCGGTGCGCGAGCGGTGCTGCTGATCAGC 5644

Db 3185 TGTCGCTGTTGATGATGCGCGCAATCATGCGCGGTGCGGATACGCGCACGCTGACCGCGC 3244

Qy 5645 CTGCGATTTAGACTGGCGTCCAGCTCACTAACTTGGCGCGCATTTGTAGAGATGTC 5704

Db 3245 CTGTGACCTTGGCTGCGCATTTGATGACCAATTTGCTGCGATATTTGAGCATGTC 3304

Qy 5705 CGAAATAGTCTGCTATCTGCGCAATCTGGCTCGATCAGCGGGATTAACGCGCGCA 5764

Db 3305 GCATCGGCGCGCTGTTATCTCGCGCAAGCTGGCTGGAGCATGAAGTCTGAACAAAGA 3364

Qy 5765 TAGCTGACTGACCGCAACATCTGTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5824

Db 3365 GAATTTATGCGGCACTGAAACCGCTGAGCGCTGAGCGGTATCGCGCGCTGCTGCTGCTGCTG 3424

Qy 5825 GGAGCGGAACCTTATCATCTCGCGCGATCGGTTTACCGGTTTACCGGCTCGCTGCTGCT 5884

Db 3425 GGAAGCAGAACCTTACTATTTGCTGCGCACAGCGCGCTGGCAGGTTGCGCTGCTGCTGCTGCT 3484

Qy 5885 GCGTGGGCGCATCGCTACGCTGCGCGGTTTATCGGAAATGCGGCTCAAAATTTGACGA 5944

Db 3485 CGCCTGGGCAATCGCTACGCGCAAGCAGGTTTACCGGAAATAGGTGTCAAAAGTTGAACA 3544

Qy 5945 CGCCGCTGTGACCGCTGGATTCAGCGAGCGCACCAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 6004

Db 3545 GGCCGCTGACGAAGCTGGGATCAGCGGAGTCAACGACCGCGCGCAAAATTTAACGCT 3604

Qy 6005 GCTGTGTGAAGGGGAGGTTTGGCGATCACTTTCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6064

Db 3605 GCTGTGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3664

Qy 6065 GCGTGTCTGTGGCAGCTCTCTGTTGATTTTACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6124

Db 3665 TCGCATCTCTGCGAGCGCGC-----TCTAGCGCCATGCTTTTCCGAGCGCTGCTGCTGCTGCTGCT 3718

Qy 6125 TCGAGCTTATTCAGCGGTGGCGGTAGAGGAAACCAAGCAGCAGCGAGCTTTCAGCGCG 6184

Db 3719 TGAAGTTTTCAGCGGCGCGCATAGAGGAAACCAAGGAAACCAACCTTCTTTTGGCC 3778

Qy 6185 CGCAGCGATGATGATGCGGTGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6244

Db 3779 CTGACGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3838

Qy 6245 ATATAGCGGAACGCGCAGCGTTGATGCAACAGCGCATCTGTCACCATGAAAGTAGAGCGCG 6304

Db 3839 ATATAGCGGAATGGCAACGTTGATGCAACAGCGCATCTGTCACCATGAAATAGAGTAAT 3898

Qy 6305 CGGTACGTCGTCATTCGCGCACCAATCCACTGCAAGCGCGCATCTGTCACCATGCAACCGCA 6364

Db 3899 CCATAGCGCGTACCTGCGCAATCCACTGCAAGCGCGCATCTGTCACCATGCAACCGCA 3958

Qy 6365 TAAATCAGCAAAATCGCCAGTACCGCAACACCGCATGAAAGATCGTTGAGCTCAAC 6424

Db 3959 TAAATCAGCAAAATCGCAATGATGCAACCAACCGCATGAAAGATCGTTGAGCTCAAC 4018

Qy 6425 TTACCGTGTGCGGTTTATGTCGACAGATGTCAGCGCGCATCTGTCACCATGCAACCGCGTGTG 6484

Db 4019 GCACCTTTACGCGGTTTATGTCGAAAGATGTCATCTGTCACCATGCAACCGCGTGTG 4078

Qy 6485 TATTATGCGACAGCGCGCTTACGATTTCCATCACCACCGGTTGCAACCAAGATAAGC 6544

Db 4079 TATTGTCGCGAGTGCAGCAATCTTCCATGCAATCAGGTAAACCAAGATGATCAGG 4138

Qy 6545 ACCTTCCATAACCGAGCATTTGTCCTCATTTGTCGAAAGGGAAGTA 6593

Db 4139 GCATTCGAATCCCAACATATTTCTCGGTAGAGAGCTCTGGCAGCA 4187

RESULT 8

LOCUS BX571873 349107 bp DNA linear BCT 26-SEP-2003

DEFINITION Photorhabdus luminescens subsp. laumondii T101 complete genome;

ACCESSION BX571873 BX470251

VERSION BX571873.1 GI:36787441

KEYWORDS complete genome.

SOURCE Photorhabdus luminescens subsp. laumondii T101

ORGANISM Photorhabdus luminescens subsp. laumondii T101

REFERENCE 1 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Photorhabdus.

2 Duchaud,E., Rusniok,C., Frangeul,L., Buchrieser,C., Taurit,S., Bocs,S., Bouraux-Eude,C., Chandler,M., Dassa,E., Deroose,R., Derzelle,S., Freyssinet,G., Gaudriault,S., Givaudan,A., Glaser,P., Medigne,C., Lanols,A., Powell,K., Siguier,P., Wingate,V., Zouine,M., Boemare,N., Danchin,A. and Kunst,F. Complete genome sequence of the entomopathogenic bacterium Photorhabdus luminescens Nat. Biotechnol. 11 (1) (2003) In press

JOURNAL Direct Submission

AUTHORS Duchaud,E., Frangeul,L., Rusniok,C. and Kunst,F.

TITLE Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25 rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail: lfrangeul@pasteur.fr, fkunst@pasteur.fr

FEATURES

source

1. 349107

/organism="Photorhabdus luminescens subsp. laumondii T101"

/mol_type="genomic DNA"

/strain="T101"

/db_xref="taxon:243265"

/complement(80..3225)

/gene="tccc1"

/locus_tag="plu4167"

/function="unknown, probable insecticidal toxin"

/note="Highly similar to insecticidal toxin complex protein Tccc of W14."

/codon_start=1

/transl_table=1

/product="Insecticidal toxin complex protein Tccc1"

/protein_id="CAE15539.1"

/db_xref="GI:36787442"


```

/translacion="MSTPTDALTQTPTVSVLDNRGLSIRDIFHRVWVIGGIDTRVT
RHQDARGLNHSIDRLYDAKQADNSVKPNFVQWYDIAGHALRTESVDAGRTTALND
TEGRVMTNATGVQTRRREGNTLPGRLLSVSEQVPODETQVTERFIWAENTTAEK
EYNLSGLCIRHYDAGVRLMSQSLAGAILSQSHQMLAEGQEAWSGDETVWQMLA
SEVYTTQSTTNAGALLTQDAKNQIRLADIAGQLKGSWLTVKGREQVIVKSLSW
SAAHGKREENGNGVITEYSPEPEQRLMGITTRRAEESQSEARVLQDLREYEPVGN
VISHNDAEATFRNRQKVEPNVYDLSYLQMSATGREMANIQOOSNQLPSPAMPY
PADNTSYNRYQTSYDRGGNVLQIRHSQSPATQNSYTTMTVSSRNSRNLVLSLTTPD
TQDGLPDAGHGOKTLIPGONLDMNMGELQVTPVPSRENSDSSEYYSYSGDVRLLK
VSEQGTGNSQVORVTVLPGLRLITGVADKTQNLQVITVGEAGRAQVRLVHESGK
PADIDNNVRYSYNLSSQIELDSEQILSQEYYPYGGTAAWAAKQNEASYKFI
RYSGERDATGLYXYRYPQWVGRLWLSADPAGTVGLNLYRMVRNPNITLTDHGL
APSNERNRTWFASFLPRKPDDEGMSMRQKIGRAIAGGLAIGLAAIAGTAGA
ALPVLTVGAAYGAGTALMGVNSLLEKGGALLARLVQGSKTLVQSQAAGAASSTL
AAYGARAQGVASAGAATGAVGSWINNADRIGGAIAGSAGVGTIDTMLGTSSTLT
HEVAAGAAGAGMTGTQGSFRAIGHAGIGTYTGSWIFGFLDVASNPAGHLANAVG
YAAGLGAEMAINRVICGGFLGRLGRLLGRAVSPYAGLARQLVHFSRVPPEFIPSVLGG
LAGGIGTGLHRVMGDSWWSRGLSAGSGIDHLAGMIGNIGRVRVLTITTTGIANALDYG
TSAIGAARRVLS"
complement (3220..3225)
RBS
/gene="tccc1"
/locus_tag="plu4167"
complement (3345..8051)
gene
/gene="tcccB1"
/locus_tag="plu4168"
complement (3345..8039)
CDS
/gene="tcccB1"
/locus_tag="plu4168"
/function="Unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex
protein TccB of W14 Putative transmembrane protein."
/codon_start=1
/transl_table=11
/product="Insecticidal toxin complex protein TccB1"
/protein_id="CAE16540.1"
/db_xref="GI:36787443"
/translacion="MSTMXEQNESQSDALVTGYMNFVASTLKGVDGQPVTVDELVE
YLLIDPEVADEVETSEARAQIASIOQYMTRLVNGSEPGQRAEMLSTANENWRDNSQYA
IWAAGAEVNAETNIAISITRQEKSHFSELETNLNQLRDPDQVDAVLGVLENEFA
VSNLVLISGYINODKFDQAIYFIFGRTTTPKPYRYVWQMDLSKNRQDPAGNPVPCW
NDMORILTPLSGDTVLEHTVPFVNDRLYVAVWERDPAVOKDAGKNIGKTHAYSIK
FGYKXDDITWAPNTTILMTQOADESSQTSRSSLIDESKTLRKYNLLATDPSIDP
TEETDSNPGRLMUGVVRNFVTGPHQNPAPVGYLYCDSAFNHHILRPLSKDFLST
YRDETNDQCLQFAVDYKXVITKISGKPTEDPENTGWGVDDLLKLTGTGAYYID
QDGLTLHTQTTNGSFVNRHIFGYNDIVIDSKSGYGFVWGSNEGFLYDHDGNYTFH
NAAIYYPAGYGGSPVNGTWALQORVNEGWAISPLLDAAHTITVQGSYIAWEGETPA
SYLPIPGTMLLWSDKINFAISLNKLESVFTSPDWPTLTITKNFSKIADNRKPYOE
INAEVADGRNLFKHYSEQTFQLASGATYPTTYTISEADPSTDPDKNYLOVCLNVWDH
YDRSGRSKWSKFNVYVELQDSKAPDAPRLISKIDSKGLVQLVDFWTSLSL
AKRNLNTFFVRLIEKANGLDSDLDYTLQADPSLEADLVTDGKSEPMDFDGSNGLYP
WELFPLPFLVATREFANEQFSPAQKSLHYIPDPAMKKNPKNAPAYMWRPLVEVHS
LSRHUGDPIDPDQAYAPHPVIYQKAVFIAYVSNLTAQGDWYRQTLRDLTQARVYVN
LAAELIGRPDPVLSSTWTPQTLTAAQOKTALRDPFHOLANSNTVLPALPGNNISH
LKLANGVFNEPLNVMLSHWDTDLARLYNLHNLTVDPGKPLSLPLYAAPADPVALLA
QRAQSGTLANGSGAILTPPYRFSAMLPRAYSAVGTILTSFGNLLSLLESSEARQGE
ELAQOQLDMSSYATLQOQALDGLAARDLALSAQTAQQRHDHYTLYQNNTSSAE
CLVMDTQSAQSLISSSTGVQTAGALKVIFNIFGLADGSGRYEGVETAIAGLMAAG
QATSVVAERLATENRRRREWEQIQOAOSEVALQQLDALAREKAATSLQQA
KAQOQIRMTLTYLTRFTQATLYOWLSQLSALYYQAYDAVSLCLSAQACWOYELG
DYATFIQGTWHDHVRGLOVGETLQNLLOMEAAVLYRHEERLNVITVSKSLGD
DGFGLKTEGKVDPLSEKLPDNYTDFGHYLRQIKTVSVTLFTLIGPYQNRATLTQTS
SSTLLTADINGVKRLDPTFGEGDATHIVTNRSQQVALLSGGINDAGSFELRLEDER
YLSFEGTGAWSKWTLTTPFRSVDHIDDKTLKADENQQAALLANMDDVLVQVHYTACDGG
ASFANQVKKTLF"
complement (8046..8051)
RBS
/gene="tcccB1"
/locus_tag="plu4168"
complement (8134..11034)
gene
/gene="tcccA1"
/locus_tag="plu4169"
complement (8134..11034)
CDS
/gene="tcccA1"
/locus_tag="plu4169"

```

```

/function="Unknown, probable insecticidal toxin"
/note="Highly similar to insecticidal toxin complex
protein TccA of W14."
/codon_start=1
/transl_table=11
/product="Insecticidal toxin complex protein TccA1"
/protein_id="CAE16541.1"
/db_xref="GI:36787444"
/translacion="MNQALASPLISRTBEIHNLPGKLTDLGYTSFVDFVVRMPREFIRE
HRADLGSAAKMYDLAVGASHOVLHFRNRSLSEAVQFSRLSPFSDPVDYANQFLDA
NTWKDXAPSGSPENDAPVAYLTHIQALDQKSGASALMNTLVERRPDI GALLIN
DAKNEVTPQQLVNEILSKKLSUTLEAENVAKLSSTTRPNNLPHYHQIQIQT
QASVLEITQDITLQPLDLPQNFQWAKGKLSATAGALRLOIMASLPFHPQEQIKM
VETVQGFYQNLNGDSLLTANSFSDMTVLTDRSLTVPQVELMLCSTVGGSTVVKSDN
VSSGDTTAPFAYGARFIHAGKPOAITLSRSGAEAFALTNNLTDDKLDINRTVRL
QKWLNPYEDIDLLVTSAMDATCNTALSMNDNTLRMLGVKPKYQAKYGVAKFAGW
LRVTPPAIIPATFPAQOVNSVCTFPTFVINDQDFVYTLTTCGDGAVRKHISTALG
LNHQFLLADNATIRQOQNTQSTLNCNLFVVSAPYRLANLARILGTNPESFCALVDR
LDAGTGVVQQLAGKPIIMVQKDSPLAADLSLLQALSVIAQWQQRDLPIETLLLS
LSSGSTTAQGTDDQLNFIQVQWQLSFTVSATLSRSGAPLVDTNGHAI DWFALLS
AGNSPLIDQVGLVTDAGTESVITMVKTSLSLSDKDLKIAITLNTLYAQOQTQGVV
VSLIAQTLNVQSLSQPLALLRWSGOTTYQWLSATWALKDTVKIAADI PAGYLSLLREV
RBSLITQOFTLSPAWQTLDDYPAYFGASAEVTDISLMMLYTHSRYSLLDLQWKAG
GCTEDDVIALYLRVANAATPLSQADAQTLATLLGWETNELQAAHEVLGGIAKTFPQLD
TLRLQQAQSGTGLGVTQQQGYVLSRSDSTLLWQSTQQAQVAVGSHVKGSH"
complement (11462..11943)
/locus_tag="plu4170"
complement (11462..11929)
/locus_tag="plu4170"
/function="Unknown"
/note="unnamed protein product; Similar to pyocin S3
immunity protein and to Photorhabdus luminescens proteins"
/codon_start=1
/transl_table=11
/protein_id="CAE16542.1"
/db_xref="GI:36787445"
/translacion="MAINIERALINSLKTYQEIFDEGLIPYKTKPGYGPDSIDSLDM
AKEGIFLFRNKVLTETLTFINHKRPNFIENKLPSPILPMSRKLIPHQFKPE
KSLPFRRLTKIEIGWTELYTLDDFRITPSMQVDYDQERVLVTLFTSEIRW"
complement (11938..11943)
/locus_tag="plu4170"
complement (12038..12505)
/locus_tag="plu4171"
complement (12038..12505)
/locus_tag="plu4171"
/function="Unknown"
/note="unnamed protein product; Similar to pyocin S3
immunity protein and to Photorhabdus luminescens proteins"
/codon_start=1
/transl_table=11
/protein_id="CAE16543.1"
/db_xref="GI:36787446"
/translacion="WTVNVDAALINSLGKSYQEIFDEGLIPYKTKPGYGPASFIALNM
AKEGMHLAFKEDGKILFAVELLDDQKPLYPQFNEPLSPKPLMTRVWVHEQFKPE
KALPFRKFKDVGWTELYTLDDFRITPSMQVDYDLELQKSVAPLSLSEVLM"
complement (12597..13081)
/locus_tag="plu4172"
complement (12597..13064)
/locus_tag="plu4172"
/function="Unknown"
/note="unnamed protein product; Similar to pyocin S3
immunity protein and to Photorhabdus luminescens proteins"
/codon_start=1
/transl_table=11
/protein_id="CAE16544.1"
/db_xref="GI:36787447"
/translacion="WIDVIEILINSLGKTADMLVENKLIPEEKFEPAFGYGRFSVEPE
TGLTLVEDIISTVLSVQPTLINTVDDGTYSQDMPAPLHKAKAIYRAFMGEEDYS
SGPERIPVIGLVGGFDAYLRKNDQYNTPEIRFLYQTLRDLRDALLFEQSVNND"
complement (13076..13081)
/locus_tag="plu4172"
complement (13174..13652)
/locus_tag="plu4173"

```


CDS

complement (13174..13635) /locus_tag="plu4173" /function="Unknown" /note="unnamed protein product; Some similarities with pyocin S3 immunity protein and similar to Photorhabdus luminescens proteins" /codon_start=1 /transl_table=11 /protein_id="CAE16545.1"									
Query Match Best Local Similarity 62.0%; Pred. No. 1.3e-278; Length 349107; Matches 2211; Conservative 0; Mismatches 1347; Indels 8; Gaps 2;									
QY	2532	AATACGATGTGATTTTGGTTCGGTCTGGACTGGCGAATGGCTTGATTCGGCTGCGCTGCG	2591						
DB	202944	ACTGGGATCTGATTTCTGGTGGGAGGAGATTGGCCAAACGGACTGATCGCAATGCGTTCC	203003						
QY	2592	GTCATTTGACGGCACAACTGAATGCGTGTTCCTGGAGAGCGATCGGATCCGGCAGGCA	2651						
DB	203004	AGCAGTGTAAACCAATCTCGAGTGTTCCTTTATTGAACAACAGAAACAATAGGAGGCA	203063						
QY	2652	ATCATACCTGGTGTTCATACAGCATCTCAGCGCGAACAACCTTCGCTGGCTGCAAC	2711						
DB	203064	ATCACAGTGGTCAATTCATCAACATGATCTTCTGAGCGGAAACATGATGGATAGCAC	203123						
QY	2712	CGCTGATTACCGTGGTGGTTCAGGTATTCAGGTGGTTCCTGCGCTGCGCCGCAATC	2771						
DB	203124	CGCTGATTACCTATCGTGGTTCAGGTACGAGCTCATTTTCCAGCATTTCAACGCACAT	203183						
QY	2772	TGGACGGGATTAATTTGTTCCATCGCATCAGGCGATTTTGGCCGCGCATCTTTACGGCGGA	2831						
DB	203184	TGCCAATTCATATTTTCAGTATTCATCCAAACACTTTTGCAAGCATATCTCCATGCTATP	203243						
QY	2832	TGGGTGACGATCTGTGGACAAACACAGCGGTACACAGTAAACCCAGCGAGTGACGC	2891						
DB	203244	TGGGCGAAGTATACAGATCGTTTATTTGGTACAGAGCTGACTCCACAGAAAGTTTACT	203303						
QY	2892	TGGCGGATGGCGGTGAATCTGTGGCGAAGTGGTGAATGATGGTGGCGGCTGCGAGCGGA	2951						
DB	203304	TACAGGACGGCTCGTCTCTAAGTGTGGCGCAGTCATTGATGGGAGAGGCTGGCGACCAG	203363						
QY	2952	CGCCACATCTGACGCTGGGTTATAGGTGTTCTTTGGACAAGATGGGAGCTGGCGCAGC	3011						
DB	203364	GACCAATTTATAGGAGTGGCCACCAGGCAATTTTGGTCAGCAATGGGAGCTGGAAGAT	203423						
QY	3012	CGCAGGCTGACAGCGCGATCTCTGTAGTGGATGCCACCGTCCATCAGCAAGCGGTATC	3071						
DB	203424	CGCACTCTTTAAACCCACCCGATTTTAAATGGATAGTGTGGGACAGGATACAGGTATC	203483						
QY	3072	GTTTTGTCTACAGCTGCGGCTCAGCGCGATCGGCTATTGATTGAAGATACCCATTACG	3131						
DB	203484	GATTTATCTATGCTGCGGTTCTCTCAACTCGTCTGCTGATAGAGNACTCATTAGC	203543						
QY	3132	TTAAACAGCCCGGCTGGCGGAGAAACACCGCTGCTGAGCACATCGCCGACTATGCAATC	3191						
DB	203544	TTGATCGGGGCGCACCTGATAAGGCTTTGTCGACGGCTACTATCGCAGATACGCGAAGA	203603						
QY	3192	AGCAAGGCTGACGCTGATACGCTGCTGCGTGAAGACACCGCATATTACCGATTACC	3251						
DB	203604	AACATGGATGGAACATGGGTAAACATATTCGAGAAGAGACGGTGTCTTCCAAATACAC	203663						
QY	3252	TGACGGGCAACATCGATCGATTCTGGCAACACAGCAGCGCGGCAAGCTGACGCGGCTGC	3311						
DB	203664	TTACGGGAGATTTTACCTCTTTCTGGGACACGCTAGCAGACAGCCACCTGTGGGTTAC	203723						
QY	3312	GGCGCGGCTGTTTATGACCAACACCGGTTATCTCTGCGCTGCGCGGCTGAGCGG	3371						
DB	203724	GTGCAGCTTTGTTTACCCCAACAAAGGCTACTCTCTGCCACACGCTATTTCGGTTGGCAG	203783						
QY	3372	AGTTGGTAGCAGCGGTGTTGCCACCGATGCCCTCAGCGCTCAGCCACATATCGAACGCT	3431						
DB	203784	ATCGTATTGTTGCTCTGCGGAGCTTACCGATACCTCTTATTCAATTACCTCAGGGATT	203843						

QY	3432	TTGCCCGTCACAGTGGCGCGAACAAGCGATTTTTCGGTCTGCTAAACCGCATGCTGTTTT	3491						
DB	203844	ACGCACGACAAACAGTGGCAACACACAGCGCTTTTTCGGTCTCTAAATCGCATGCTCTCC	203903						
QY	3492	TGGCCGGTAAAGCCGACGAGCGCTGGCGGTGATGCAACGTTTTTACCGCTCGATGCG	3551						
DB	203904	TCGCTGGGGATCCCAACAACGTTGGCAGGTAATGCAACGTTTCTATCAACTTTCCCAA	203963						
QY	3552	GGTTAATTAGCGCTTTTACCGCGGCAACTGCGCTCGCGGATAAAAACGCGGATTTCTGT	3611						
DB	203964	ATCTGATTGACGTTTTTATGGGAGCAACTTAATTCGTCGACAGGCCCGGATTTCTCA	204023						
QY	3612	GCGCAAGCCCGGTGCCCCATCGGTGAAGCGCTGCGCGCTGTGAA-----TTCTGT	3666						
DB	204024	TAGGTAAACCAACCACTGCGGATTAAGGGTGTCTAAAGGCAATGTTAAACAACAACA	204083						
QY	3667	CGAACCCAGGAAGAAAAATGAACGCCTTATGTGATTGGCGCAGGCTTTGGCGGCTG	3726						
DB	204084	AGCTTACGGGTTTTTATTGATTAAAGCGCTGGTAAATTTGGTGTGGTGTG	204143						
QY	3727	GCGCTGGCGATTTCGCTCAAGCGCGGCGCATACCAACCACTTACTCGACAGGCGGAC	3786						
DB	204144	GCACTGGCAATAAGGCTCCAGTCTGCGGGGATTCGACATGATTTTGGAGCAACGGAT	204203						
QY	3787	AAACCGGCGGACGCGCTATGTGTTGAGACAGTGGCTTTACCTTCGATCCCGACCC	3846						
DB	204204	AAACCGGTCGACCGGCTTATGTTTATAAGGAACAGGGATTCACCTTTGATCCGCGCC	204263						
QY	3847	ACGGTGATCAGGATCCAGCGCCATCGAAGTGTTCACGCTGCGCAGGAATCGCTC	3906						
DB	204264	ACCGTAATTCACCGCTCTTAATGTCAATGAAGAAATTTGTACCCAGGCTGGTAAAGT	204323						
QY	3907	AGCGATTACGTCGAGCTCATGCGGTAACGCGCTTCTATCGCTGTGCGGAAGATGCG	3966						
DB	204324	GCTGATTACGTCGATTTACTTCCCGTACACCTTTTATCGGCTCTGTTGGAGTCGGT	204383						
QY	3967	AAACAGCTTGATTACGACAATAACAGCGCTGCTGGAGCAGCAGATCGCCAGTTCAAT	4026						
DB	204384	AAGATGTTGATTACGACAATGATCAGCAACATCTGGAAGCACAATCCATACGTTCAAT	204443						
QY	4027	CGCAGAGTGTAGAAGCTATCGTCAATTTCTTGCTATTACCTGAGTATTTAGAGAG	4086						
DB	204444	CCAGAGATGTTAAACGGGTATCGACGCTTTCTAGACTATTCGCGGAAGCTTTAATGAA	204503						
QY	4087	GTTATCTGAAACTCGGACGCTGCGCTTTCTGACAGTGGCTGACATGCTGCGCTCGCG	4146						
DB	204504	GGTACCTGAAACTTGGAAACGGTGCCTTCTCTCTTTTCGCGACATGTTGAGTGGGCT	204563						
QY	4147	CGCAGTTGGGACGTTGCAAGCATGGCGCAGCGCTCTACAGCATGGTGGCGAAAATTTAT	4206						
DB	204564	CCACAGTTGATACGCTTACACGCAATGGCGTAGTGTATTAGTCAGGTGGCAGCTTATC	204623						
QY	4207	CAGGACGATCATTCGCTCAGCGGTTTTCTTCCACTCATTTCTGTTGGCGGTAATCCT	4266						
DB	204624	AAGGATGAGACCTGCGCCAGGCAATTTTCATTTCTCATCTGCTATTGATAGGAGTAAC	204683						
QY	4267	TTTGCAACGCTCATTCGATCTATACCTTAATTCATGCGCTGGAGCGTGAATGGCGGCTG	4326						
DB	204684	TTTGCTGATCTTCTATCTTACTTTGATCCAGCAGCTGGAACGAGATGGGAGTCTGG	204743						
QY	4327	TTTTCGCGCGGCGCACCGCGCGCTGTTGTCAGGCGATGGCGGATGTTTCGAGGACTTTG	4386						
DB	204744	TTTTCGCGGAGGAGAAACCAAGCGCTCTGGTTGAGGCAATGGTGAATAATTTGTTACTG	204803						
QY	4387	GGCGCGGAGCTTTACTGATGCCAAAGTACCGCAGCTGGAACCAAGCGGCAATCGCAT	4446						
DB	204804	GGTGAGAAATGAGCTTAATGCAAAAGGTAAACCGCTTCCACCCCATGCGCAATCGGTC	204863						
QY	4447	AGCGCGCTTCAGTTAGAGGCGGACGAGCTTTCGATGCGCGCTGTGGCTTCCCAATGCC	4506						
DB	204864	ACAGGTGTCATAGCTGATGGACGAATATGACATGTGAGCTTGTGGCTCAATGCC	204923						

/translation="MKKTVVIGAGPGGLALAIRLOAAGIPTVLLRQDKPGGRAYVMH
DGFPTDAGPTVITDLEALFTLAGRMEDYVRLLEPVKPIRUCWSGSKLDITAND
SABLEAQITQFNPRDVEGYRFLYQAQVFGYLRGLGSPFLSRDMLRAGPOLKL
QAWQSVYQSVFIDEHLRQAFSPHLLVGNPFTSSYITLIHALREMGWPFPEG
GTGLVNGMFLFDLGGIEELNARVELVADNRVSOVLADGRI FOTDAVASNADV
VNTYKLLGHLPVGGKRAALRERKSMNSLFLVLYFNQPHSOLAHHTICFGRPREL
IDEITGSLHADDFSLYLHSPCVTPSLAPGCSAPVYLAPVPHLGNAPLDWAORPK
LDRIPDYLEERYMGLRSQLVQRIPTPADPHDILDALHLSAFSIBELLTQSARFP
HNRSDIANLYLVGAGTHPGAGIPGVVASAKATASLIMIEDUQ"
1486..2415
/gene="crtB"
1486..2415
/gene="crtB"
/codon_start=1
/transl_table=11
/product="prephytone pyrophosphate synthase"
/protein_id="AAA24821.1"
/db_xref="GI:148403"
/translation="MSQPLLHATQTMANGSKSPATAAKLPDPATRSVLMLYTWCR
HCDVDIDQTHGFSEAAABEEATQRLRLTLTAAFEAGEMQDPAPAAFOEVALTH
GITPRMLDHLGPFAMDAQATRYVFTEDLTRYCYHVAGVGLMVARVMGRDRELDLR
ACDLGLAFTLNIAARDIIDDAIDRCYLPALWLODAGLTPENYAARENAALARAER
LIDAAEPYVISOAGLHLDLPPRCANAIATARSVYREIGIKVKAAGSSAWDRQHTSKG
EKIAMMAAPGVIRAKTRVTPRAGLWQRPV"

ORIGIN

Query Match 17.0%; Score 1189; DB 1; Length 2415;
Best Local Similarity 68.6%; Pred. No. 1.6e-239;
Matches 1655; Conservative 0; Mismatches 755; Indels 3; Gaps 1;
QY 3678 AGAAAAATGAAACGACATTATGTGATGGCGAGGCTTTGGCGGCTGGCGCGAT 3737
DB 4 AATAAGGATGAAAAAACCGTTGTGATTTGGCGAGGCTTTGGTGGCTGGCGCTGGCGGAT 63
QY 3738 TCGCCTGCAAGCGGGGATACCAACACCTTACTCGAGCAGCGGACAAACCGGGCG 3797
DB 64 TCGCCTGCAAGCGGGGATGAGGATCCCAACCGTACTGCTGGAGCAGCGGACAAACCGGGCGG 123
QY 3798 ACGCGCTATGTGTTTGGAGCAGTGGCTTTACCTTCGATGCGGACCCACCGGTGATCAC 3857
DB 124 TCGGGCTACGCTGCGATGACACAGGCTTTACCTTTGACGCGGGCGGCGATGATCAC 183
QY 3858 CGATCCAGCGGATCGAAGAGTGTTCACGCTGGCAGGAAATCGCTCAGCGATTACGT 3917
DB 184 CGATCCACCGGCTTGAGCGCTGTTCACCCCTGGCGGCGGCGATGAGGATTTACGT 243
QY 3918 CGAGCTGTGCGGTAACCCCTTCTATCGCTGTGCTGGGAGATGGAACAGCTTGA 3977
DB 244 CAGGCTGTGCGGTAACCCCTTCTACCGACTCTGCTGGGAGTCCGGAGAGACCCCTGA 303
QY 3978 TTACGACAAATACAGCCGCTGTGAGCAGCAGATCGCCAGTTCATCCGCAAGATGT 4037
DB 304 CTATGCTAACGACAGCGCGAGCTTTGAGCGCGAGATTACCCAGTTCAACCCCGGACGT 363
QY 4038 AGAAGGCTATCGTCAATTTCTTCCCTTATTCAGTGAAGTATTTAGAGGGTTATCTGAA 4097
DB 364 CAGAGGCTACCGGCGCTTCTGCGCTTACTCCAGCGGTTATTCAGAGGAGATTTTGG 423
QY 4098 ACTCGGACGCTGCGGTTTCTGACGCTGCTGACATGCTGCGGCTGCGCGCGGCTGGG 4157
DB 424 CTTGCGCAGCGTGGCGTTCCTCTCTTTTCGCGACATGCTGCGCGCGGCGGCGGAGCTGCT 483
QY 4158 AGCTCTGCAAGATGCGCAGCGCTTACAGCATGCTGGCGGAAATTTATTCAGGACGATCA 4217
DB 484 TAAGCTCCAGGCGTGGCAGGCGTCTACAGTGGTTTCGCGCTTTATTTAGAGATGAGCA 543
QY 4218 TCTGCTCAGGCGTTTCTTCCACTCATTTGCTGCTGGCGGTAATCTCTTTTGAACGTC 4277
DB 544 TCTGCGGACGCGCTTCTCGTTCACCTCCCTGCTGGTAGGCGCAACCCCTTCAACACCTC 603
QY 4278 ATCGATCTATACCTTAATTCATGCGCTGAGCGTGAATGGGCGGTGGTTTCCGCGCG 4337

DB 604 GTCCATCTACACCTGTATCCAGCCCTTTGAGCGGAGTGGGGGTCTGGTTCCCTGAGGG 663
QY 4338 CGGCAACCGCGCGCTGGTGCAAGGCAATGCGCGCATGTGTCGAGGACTTGGGCGGCGAGCT 4397
DB 664 CGGCAACCGGGCGCTGGTGAACCGCATGCTGAAGCTGTTTACCAGTCTGGGCGGGGAGAT 723
QY 4398 GTTACTGATCGGAAGTGAAGCCAGCTCGAAGCAGCGCAATCGCATTTAGCGGCGTTCA 4457
DB 724 CGAATCTCAACCGCCCGGTCGAAGAGCTGGTGGTGGCCGATAACCGGTCGAAGCAGGTCCG 783
QY 4458 GTTAGAGGCGGACGACGCTTCGATGCGCGCGCTGTGGGCTTCAATTCGCCAGCTGTGTGA 4517
DB 784 GCTGCGGATGCTCGGATCTTTGACACCGACCGCTAGCTCGAACGCTGACGTGTGTGA 843
QY 4518 TACTTACGACAAATGCTTTGCGCAACATCCGCTGGCAATGAAGCTGGGCAATTCGCTGAA 4577
DB 844 CACCTATAAAAGCTGCTCGGCCACATCCCGTGGGCGAGAGCGGCGGACGCGCTGA 903
QY 4578 GGTAAAGCGCATGAGCAACTCGCTGCTGTTGTTACTCTATTTTGGCTGAAATCAGCCGATCA 4637
DB 904 GCGCAAGAGCATGAGCAACTCGCTGTTGTGCTTACTTTGCGGCTTGAACACGCTCATTT 963
QY 4638 ACAGCTCGCGACCAACACCGCTGCTGTTTGGCGCGGCTTATCGTGAAGTTGATCGATGAGAT 4697
DB 964 CAGCTGGCGCACCATACCATCTGTTTGGTCCCCTACCGGAGCTGATCGACGAGAT 1023
QY 4698 TTTCAACAGCAGCAGCTGGCAGACGATTTTTCATCTTACCTGCAAGCGGCTTGCAGCAG 4757
DB 1024 CTTTACCAGCAGCGCTGGCGGATGATCTTCTCGCTACTACCTGCACTCGCCCTGCGTAC 1083
QY 4758 CGATCCGCTGCTGGCACCGCGGCTGGGCGAGCTTTTATGTTAGCGCGGTCGCGCA 4817
DB 1084 CGATCCCTCGCTCGCGCTCCCGGCTGGCGAGCTTCTACGTGCTGGCGCGGTCGCGCA 1143
QY 4818 TCTCGGACCGCTGACATCGACTGCGCAACAGGAAGCGCGCTTGGCGGATCGAATTTT 4877
DB 1144 TCTTGGCAACCGCGCTGGACTGGCGGAGAGGGGCGAAGCTGCGGCAACGCGATCTT 1203
QY 4878 TGCTTATCTGGAGCAGCACTACATCCCGGATTAAGTCAAGCAATTTAGTGACACAGAT 4937
DB 1204 TGACTACCTTGAAGAGCGCTATATGCCCGGCTGCGTAGCAGCTGGTGAACCGGAT 1263
QY 4938 GTTTACGCGCTTTGATTTTTCGCAACGCTGATGCCCATCACGCTCGGCGCTTTTCGCT 4997
DB 1264 CTTTACCCTCGGAGACTTTCACGACACGCTGGATCGCATCTTGGGATCGGCTTCTCCAT 1323
QY 4998 GGAGCGGATTTTACGCAAGCGCTGTTTCCGCGCGATAACCGGATCGCGATATCAG 5057
DB 1324 CGAGCGCTGCTGACCCAAAGCGCTGTTCCGCGCGACACACCGGACAGCGATTCG 1383
QY 5058 CAATCTTATCTGGTGGTGGCTACGATCCAGCGCGGCGCTGCGCGGCTGATCGG 5117
DB 1384 CAACCTTACCTGGTGGCGGAGTACTCACCTGGGCGGCGATTCCTGGCGTAGTGGC 1443
QY 5118 TTCGCGCAAGGCCACCGCAGGCTGATGCTGAGAGATCGGCGGATGATGATCGACAGCT 5177
DB 1444 CTCGCGAAAGCCACCGCGAGCTGATGATGAGGATC---TGCAATGAGCCAAACCGCG 1500
QY 5178 TTACTTGAAGTAGTAACGCAACCATGGCGTGGCTCGAAGAGTTTTCGCCACCGCGCG 5237
DB 1501 CTGCTTGACCGCCACCGAGACCATGGCCAAAGCTCGAAGAGTTTTCGCCACCGCTGG 1560
QY 5238 AAGCTGTTTGTATGACCGACCGCGGCGGAGAGCTGATGCTGATGCTGCTGCTGCTGCT 5297
DB 1561 AAGCTGTTTCGACCCCGCCACCGCGCTAGCGTGTGATGCTCTACACCTGGTGGCGCCAC 1620
QY 5298 TCGGATGATGATGATGATGGGCAACCGTGGCGGAGAGCGGCGGACGAGCATGCCGTCGA 5357
DB 1621 TCGGATGATGATGATGATGACACCGACCGCTTTCGCGAGGAGCGCGCGGAGGAG 1680
QY 5358 GAGCGCAGGCGATGATGAGCATCTGCAAAATTTGAAACCCCGCGGCGCTTACAGCGCGG 5417
DB 1681 GAGGCAACCGCGCTGGCGGCTGCGACCGCTGCGCGGCTTTGAGGGGCG 1740

Db 293 AGTGGTCTGCCACAGCGATCTCACTGCGCGCAGCGCGCTTCTCGCTCCCTTC 352
Qy 2717 ATTACGCTGGTGTGTCAGGTATCAGGTGCGGTTTTCGCGCTGCGCGCAATCTGGAC 2776
Db 353 CAGAGCTACTCTGGGCGCGGACAGGGGTGCATTTTCCCGCTTTTTCGCGCACTGCAAA 412
Qy 2777 GGGGATATTGTTCCATCCATCAGCGATTTTGGCCGCCATCTTTACGCGCGATGGGT 2836
Db 413 GCGGGCTAGCCACCAATTTCTCCGAGCGCATGCGCGGATGAATGAGCGCTGTGC 472
Qy 2837 GACGATCTGTGACAAACACAGCCGTACCAAGGTAAACCCACGACGAGTGACGCTGGCG 2896
Db 473 GCGCCATCCGACCAATGCCCGGTGCGCCATGTGGCGCCGACCACTGTGTGTGGAG 532
Qy 2897 GATGGCGGTGAATTGCTCGCAAGTGTGATGTATGTGTGCGCGCTGACGCGACCCA 2956
Db 533 GCGCGGAGCGCATGACGCGCGCGGTGTGTGTGAGCGGCGGCGCCACTCGCCTCGCG 592
Qy 2957 CATCTGACGCTGGTATCAGGTGTTTCTTGGACAAGTGGCAGCTGGCGCGACCGCAC 3016
Db 593 CATCTGACCTCGGCTACGACCTTCTCTGGGCGAGAACTGCGTATGTCTCGGCCCAT 652
Qy 3017 GGCCTGACGACCGCATCTGATGATGCCACCGTCGATCAGCAAGCGGTTATCGTTT 3076
Db 653 GGCCTACCGCCCCATCATGATGCGCGGTGAGCAGCTGGCGGTTACGCTTC 712
Qy 3077 GTCTACGCTGCGCTGAGCGCGATCGGCTATTGATGAAGATACCATTAAGTTAAC 3136
Db 713 GTCTATGTCTGCGCTGACGATGACGCTGCTGTGTGGAAGACACCTATTATGCGGAC 772
Qy 3137 CAGCGCGCTGGCGAGAACCGCTCGTCAGCATCGCGCATTTCCGCGCATTTCCGCGCGCAG 832
Db 773 GGCCTGATCTTCCCGCGCATGCTTTCGCGCGCGCATTTCCGCGCATTTCCGCGCGCAG 832
Qy 3197 GCGTGACGCTGATGACGCTGCTGCTGAAGACGCGCATATTACCGATTACCTCTGAGC 3256
Db 833 GCGTGCGCGTGAATTATGTTGTGCGGAGGAGGAGCGCATTTGCCGATAGGTTGGC 892
Qy 3257 GGAACATCGATGATTTGGCAACAGCAGC---GCGGCAAGCGTGAAGCGGCTGCGC 3313
Db 893 GCGACATCAAGCCCTTCTTCCGAGAGCGCTTTCGCGCGCTGCGCGCGCGGCTGCGG 952
Qy 3314 GCGGCGCTTTTCATGCCACCGGTACTCTTTCGCGCTGCGCGCGGCTGCGGAG 3373
Db 953 CAGGCGCTTTCACCCACCGCTACGCTGCGCGCGCATGCGCGCTGCGCGCGC 1012
Qy 3374 TTGGTAGCAGCGCTTTCGCCACCGATGCCCTCAGCTCAGCAACATATCGAACCTTT 3433
Db 1013 TCGGTGAGCGCTTTCGCCACCTTTCGCGCGCGCGCTTTCGCGCGCAGTGGCGCCAT 1072
Qy 3434 GCGCGTACAGTGGCGGAAACAGCATTTTTCGCTGCTAAACCGCATGTGTTTTG 3493
Db 1073 GCGGCGCGGTGGAACGCGCGCGCTTCTTCGCGCTGCTCAACCGCATGTGTTCCGC 1132
Qy 3494 GCGGTAAGCGCGCAGCGCTGGCGGTGATGCAACGTTTTCGCGCTGATGCGCGG 3553
Db 1133 GTTGGCGATCCGAGCGCGCTATGCGATTTCTCAGCGCTTCTACGCGCTGTCGGAAGAC 1192
Qy 3554 TTAAATAGCGCTTTTACGCGGCAACTGCGCTGCGGATAAACCGGATTTCTGTGC 3613
Db 1193 CTGATCGCGGTTTCTACGCGATCGCTCACCTTTCGCGACAGCGCGCATCTTTTCG 1252
Qy 3614 GGAAGCGCGGTGCCATTCGCTGAAGCGCTGCGCGCTGTGTAATTTCTGTGCAACCA 3673
Db 1253 GCGCGCCACCGGTATCGGTGTTTCAGGCGCTCTCTGCGCTTGTGAGACGAAGCGCA 1312
Qy 3674 GGAAGAAAATGAA-----CGCACTTATGTATGGCG 3709
Db 1313 CCGGTTTCCCATGATGCTGATGCAACCAAGACCATCCGCGCGCGCTCATCGGT 1372
Qy 3710 CAGGCTTTGCGCGCTGCGGCTGCGGATTCGCTGCAAGCGCGGCGCATACCAACCACT 3769
Db 1373 CCGGCTTTCGCGGCTCTCTCTCGCATTCGCGCTGAGCGGCGGCGCATCCGACACCG 1432

Qy 3770 TACTGAGCAGCGCGCAAAACCGGCGGACGCGCCTATGTGTTTGAAGACAGTGGCTTTA 3829
Db 1433 TGTTCGAGCAGCGGACAAAGCCCGCGCGCTACGCTTATGAGCAGGACGCGTTCA 1492
Qy 3830 CTTTCGATGCGGACCCACGCGTATCAGCATCCAGCGCCATCGAAGAGTTGTTCAACGC 3889
Db 1493 CTTTCGATGCGGCGCCACCGCTATCAGCATCCACCTGCTGGAAGAGTTGTTCCCG 1552
Qy 3890 TGGCAGGAAAATTCGCTCAGCGATTACGTCGAGCTGATGCCGTAAAGCCCTTCTATCGCC 3949
Db 1553 CCGCGCGCGCAAGCTGTCGAGCTATGTGAGCTGATGCCGTTCGCGCTTCTACCGGC 1612
Qy 3950 TGTGCTGGAAGATGGCAAAACAGCTTGATTACGAATAATCAGCGCTCTGGAGCAGC 4009
Db 1613 TGTCTGCGCGACGCGCGCGAGTTTCGACTATGTGAACACACGCGCGCTGGACGCC 1672
Qy 4010 AGATCGCACGTTCAATCCGCAAGATGTAGAAGGTATCGTCAATTTCTTGTGCTATTTCAC 4069
Db 1673 AGATCGCACCTTCGATTAAGCGGACGTGGAGGCTATCGCGCTTCTCGCTATTTCG 1732
Qy 4070 GTGAAGTATTTAGAGAGGTTATCTGAACCTCGCACGCTGCGGTTTCTGCAAGTCCGTG 4129
Db 1733 AGCGGCTTTCGAGGAGGTTATCTCAAGCTCGCGCGGTGCGCTTCTCGAATTTCTCCA 1792
Qy 4130 ACATGCTGCGCTCGCGCGAGTTGGAGCTGTGAAGCATGGCGCAGCGCTTACAGCA 4189
Db 1793 GCATGATGAAGCGCGCGCCAAAGCTGGTAGGCTGGAAGCTTGGCGCTCGCTTATGCGCA 1852
Qy 4190 TGTGCGCGAAATTTATTCAAGACGATCATCTCGCTCAGCGGTTTCTTCCATCTCATTTGC 4249
Db 1853 TGTCTCGGCTTCACTCGGACGATCACCTCGCGAGGCTTCTGCTTCAATTCCTCTG 1912
Qy 4250 TGTGCGCGGTTAATCTTTTGAACGCTCATCGATCTATACCTTAATTTCAATGCGCTGAGC 4309
Db 1913 TGTGCGGCGAATCCCTTCTCCACTGCTCATCTATGCCCTCATCTCATGCTGCGTGAGC 1972
Qy 4310 GTGAATGGGCGTGTGTTTTCGCGCGGCGGACCGCGCGCTGTGTGAGGCGATGGCGC 4369
Db 1973 GCAAGTGGGCGGTGTTTCTTCCCGCGGGGAAACCGCGCGCTTGTGCGCGCATGGTCA 2032
Qy 4370 GACTGTTGAGGACTTGGCGCGGAGCTGTATTGAAATGCGCAAGTGAGCCAGCTGGAAA 4429
Db 2033 AGCTGTTTACCGACCTCGCGCGCGAGATCCGCTCACAGCCCGGTGAGCAGATCGTGG 2092
Qy 4430 CCAGCGGCAATCGATTAGCGCGCTTCAAGTTAGAGCGGCGGACGCTTCGATGCGCGCGC 4489
Db 2093 TGGAGGCGCAGCGCGCACCGCGGTGAAGCTCAAGAGCGCGAGACCTTCCGCTTCGACC 2152
Qy 4490 CTGTGCGCTTCAATGCGGACGTGTGTGATACCTACGAATACTGCTTCGCGCACCATCCGC 4549
Db 2153 TCGTGCCTTCCAATGCGGACGTGTGTGACACCTATCGCCACCTTCTGCGCGGCGCGCGC 2212
Qy 4550 TGGCAATGAACGTCGACATCGCTGAAGCGTAAAGCGATGAGCAACTCGCTGTTTGTATC 4609
Db 2213 GCGCTGCTCGGAAGCGCGCGCTGCGCAAGACGCGCATCTCATGTGCTGTTGCTCA 2272
Qy 4610 TCTATTTTGGCTGAAATCAGCGCGCATGAACAGCTCGCGCACCAACCGCTGTGTTTGGCC 4669
Db 2273 CCTATTTTGGGCGCGGCGCACCTGGGACCATCTCCAGCACCAACAGCGTGTGTTGCTC 2332
Qy 4670 CCGCTTATCGTAGTTGATGATGATTTTCAACAGCAGCAGCTGGCAGCAGATTTT 4729
Db 2333 CCGCTATTCGCGGCTCGTGTGACGATCTTCAAGGGCCGAACTTCCCGACGACTTCT 2392
Qy 4730 CACTTTTACCTGACGCGCTTCAGCAGCGATCTGCTGCGCACCGCGCTGCGGCA 4789
Db 2393 CGCTTATCTCCAGCGCGCCACCGGTGACAGAAATCGCTGCGCGCGGAGGCTGTACCG 2452
Qy 4790 GCTTTTATGTGTAGCGCGGTGCGCATCTCGGACCGCTGACATCGATGCGCAACAG 4849
Db 2453 CTTTCTACGTGCTCTCGCGGTGCGCATCTGGGCAAGCGGACATCGACTGGAGGTGG 2512

QY 4850 AAGGACCGCGTTCGCGCATCGAATTTTGTCTTATCTGAGCAGCACTACATGCGCGGAT 4909
DB 2513 AAGGCGCGCTCTATCGGACCGCATCTCGCCCATCTGAGAGCGGCTACTGCGCGGGC 2572
QY 4910 TACGTAGCAATAGTGTGACACAGAGATGTTTACGCGGTTTGATTTTCGCGACAGCTGC 4969
DB 2573 TCGCGACAGCCTCTGCTACCTCGGCATCTCTACCCCTTTCGCGTTCGCGAGCAACTGT 2632
QY 4970 ATGCCATCAGCGCTCGGCTTTTCGCTGAGCCGATTTTACGCAAGCGCTGCTTCC 5029
DB 2633 CGCGCCATCAGGCGCTTCGCTGAGAACCGCTCTCACCAGTCAGCGGTGTTC 2692
QY 5030 GCCCGATAACCGCATGCCGATATCAGCAATCTCTATCTGCTGCTGCGGTAGCATC 5089
DB 2693 GCGCGACAAACCGCATGCAAGATTGCCAATCTTATTTCCGCGGCGCGCACCATC 2752
QY 5090 CAGGCGCGGCGTTCGCGCGTGTGATCGGTTTCGCGCAAGCCACCGCCAGGCTGATGCTGG 5149
DB 2753 CGGCGCGCGCTTCGCGGCGTGTGGCTCGGCCAAGCCACCGCGGCGCTCATCTCG 2812
QY 5150 AGATCGCGCC-----GAATGAATCGACAGCCTTTACTTGTAGCAAGTAAC 5194
DB 2813 CGGATCGGCGGTGACCGCGCGATGAGCAGCGCTCGGCTACCGTCTCGGTGCCAGCA 2872
QY 5195 GCAAACTATGGCGTGGCTCAAGAGTTTCGCCACCGCGCCAAAGCTGTTTATGATGCACC 5254
DB 2873 GAGGCCATAACAGGGCTTCAAGAGCTTCGCGCGCGCGCGCTGTTCCGCGCGG 2932
QY 5255 GACGCGCGCAGCAGCCTGATGCTGTATGCTGTGTGTCTCTACTGCGATGATGATGA 5314
DB 2933 CATCGGGAAGACCGCGTGTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA 2992
QY 5315 TGGCGAAACCTGCGCGAGGCGGCGCAGCAGCATCGCTCGAAGAGCGCGCAGCGAT 5374
DB 2993 CGGCGAGGAACCTGGGCGACCGGCGTGTCTCGAGCGCGCGCTCGCGCGCGAG---CGGT 3049
QY 5375 GCAGCATCTGCAATTTGAACCCGCGCGCTACAGCGCGCGCGCACATGATGAACCGCG 5434
DB 3050 CGATGCGCTTATGAGGAGACCGCAACGCTATCGCGCGCGCGCTCGCGCCATCGCG 3109
QY 5435 GTTATGCGGCTTTCAGGAAGTGGCGATCATTCACAGCTGCGCGCAACACTGCGGTTGA 5494
DB 3110 CTTTCGCGCGCTTCGCGCGGTTGTGAAGCGCAACGACATTCGCGAGCGCTATCCCGCTCA 3169
QY 5495 TCATCTGGAAGCTTCGCTATGATGATCAGCAACGACATTTACGCGAGCTTCGATGACAC 5554
DB 3170 ACTGCTGAGGGCTTCGATGAGCGTGTGAAGCGCGCGCTAGCACACCTCGATGACAC 3229
QY 5555 GCTGCGCTTACTGCTATCAGCTCGCGCGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5614
DB 3230 GCTCAGCTATTGCTACACGCTGCGCGTGTGTGCGGCTGATGATGCGGCTCATCATGGG 3289
QY 5615 GCTGCGGAGCAGCGGTGCTCATCAGCTGCGGATTTAGGACTGCGGTTCCAGCTCAC 5674
DB 3290 CGCGCGGGATGAAGTGTGTGACCGGCGGAGCGACCTCGGCTCGGCTTCAGCTCAC 3349
QY 5675 TAACTTTCGCGCGCATTTCTAGAAGATGCGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5734
DB 3350 CAACATCGCGCGCATGATGAGGATGCGGCGGATCGGCGGATCTACTGCGCGCGG 3409
QY 5735 CTGCGCTGATCAGCGCGGATTAACGCGCGGATACGCTGATGATGATGATGATGATGATGATGAT 5794
DB 3410 CTTTCTCGCAGAGCGCGCGTTCGCGAGGCGGCGATCGCGCGCGCGGAGCACCGCGCGG 3469
QY 5795 GCTCGCTTACTGCGCAGCGGTTTGTGCGGAGGCGGAGCCCTTATTCACCTCGCGCGG 5854
DB 3470 GGTGCGCGCGGTGTGCGCGGCTGCTGCGAGCTGCGCGGAGCCCTTATTCAGCAGGCGCT 3529
QY 5855 ATCGGCTTACCGGCTTACCGCTGCGCTCGGCGTGGCGCATCGCTACGCGCTCGCGCGG 5914
DB 3530 CATCGGATGCGCGCGCTTCGCGCGCGCGGCGGAGCGGTGGCGCGCGCGGTGGGT 3589
QY 5915 TTATCGCGAAATTGGCGCTCAAGTTTACGACCGCGGCTGTGCGACCGCGCTGGGATTCAGCGCA 5974

DB 3590 CTATCGCGCATCGCATCGAGTCTCGTAAGCGCGGCGCGGCGGATATCCCGGT 3649
QY 5975 GCGCACAGTAAGGTGAATACTGGCGCTCTGCTGTAAGAGGCGGACGTTTGGCGATCAC 6034
DB 3650 CTCACCTCCACGCGCCAGAGCGCGCTTCTCGCGGCTGCGCTCGCGAGGCGCTCGC 3709
QY 6035 TTGCGGTGTGCTCG---TCCTGAACCGCGTTCGCGCTGCTGCTGCGAGCGTCTCTCGTTG 6091
DB 3710 CACGCGTTAGAGCGCGCGCGCGCTCCCGCGCATCTTTGGACGCGTCCCGCATG 3769
QY 6092 A 6092
DB 3770 A 3770

RESULT 11
AF218415
LOCUS
DEFINITION
Bradyrhizobium sp. ORS278 DNA linear BCT 22-JUN-2000
cyclase (crtY), phytoene desaturase (crtI), phytoene synthase
(crtB), and beta-carotene ketolase (crtW) genes, complete cds.
ACCESSION
AF218415
VERSION
AF218415.1 GI:8650414
KEYWORDS
Bradyrhizobium sp. ORS278
SOURCE
Bradyrhizobium sp. ORS278
ORGANISM
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Bradyrhizobiaceae; Bradyrhizobium.
REFERENCE
1 (bases 1 to 6635)
AUTHORS
Hannibal, L., Lorquin, J., D'Ortoli, N.A., Garcia, N., Chaintreuil, C.,
Masson-Boivin, C., Dreyfus, B. and Giraud, E.
TITLE
Isolation and characterization of carboxanthin biosynthesis genes
from the photosynthetic bacterium Bradyrhizobium sp. strain ORS278
J. Bacteriol. 182 (13), 3850-3853 (2000)
JOURNAL
20309720
MEDLINE
10851005
REFERENCE
2 (bases 1 to 6635)
AUTHORS
Giraud, E. and Angles d'Ortoli, N.
TITLE
Direct Submission
JOURNAL
Submitted (02-SEP-1999) Laboratoire des Symbioses Tropicales et
Mediterraneenes, TA/10J, Campus de Baillarguet, Montpellier 34398,
France
FEATURES
source
1..6635
/organism="Bradyrhizobium sp. ORS278"
/mol_type="genomic DNA"
/strain="ORS278"
/db_xref="taxon:114615"
complement(273..1223)
/gene="crtB"
complement(273..1223)
/gene="crtE"
/codon_start=1
/transl_table=11
/product="Geranylgeranyl synthase"
/protein_id="AAF78199.1"
/db_xref="GI:8650415"
translation="MHKPVLDLTDAAFTQDRWGRIGIEAVAEAMAFGTTVPAPLQA
GMSHAVLAGGKRYGMLVGLSGDLGVEQLLSAVAIETTHAASLVVDLPCMDMDA
RRRSQPAHVAFGEATLASSIALIARAMEVARDROLSPASRSSIVDTLSHAIQPO
ALCGQYDDLPFPYATQDLHRYQKTSALFVAARCPALAEVDPTLLRIARAG
QLGVAFOIFDLDLDTGDAHAGKDVQGDHGTVTATLLGPARAAERAADELAQVK
ELRETVGFRALDLIRMAARTAGTGGKKSAGRDDLRPHAG"
1446..2630
/gene="crtY"
1446..2630
/gene="crtY"
/codon_start=1
/transl_table=11
/product="lycopene cyclase"
/protein_id="AAF78200.1"
/db_xref="GI:8650416"

```

/translation="MSRDADVIVIGGLAGCLIALRLTDARPDRLRVIIIEGASAGN
HTWFFGDISDQAHWGLRVLGHRWPGVEVFAEHAIRKLSTAYLSMTSTRURAEVQ
RPERILRDATAIAFCADVDGTRFRPSYERLLBEGYVKLGHVFPFPPFRSMVRV
LAAPGLDVPVIMYDATTVAQSDYRFVYTLPLDQRLIEDTYSYDGLPEQLVHORI
ARYALAKQWIAEIIITRAECQVLPVILAGPSGLVSKPSPRVRVLAALLVHETTCYSL
PDARVADLLTARLAQRGALSADARETIDYGRITWRERGYYRFLNRMFLKAAEPSR
RHRILLARFVGLDQALIERFYARIQPDQKLRVFMHMLKPPPIPSALACLPEASAFK
TP"
2627. .4147
/gene="crt1"
CDS
2627. .4147
/gene="crt1"
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="AAF78201.1"
/db_xref="GI:8650417"
/translation="MSQEMONAKTAVVIGSGFGGLSLAIRLQSGAGIATTLVEKRDKFG
GRAYVEQDGFDFDAGPTVITDPTCLQELFALSGRKLENYVELMPVSFYLQRLWEDGA
TFDYVNDQALERQIAAFCADVDGTRFRPSYERLLBEGYVKLGHVFPFPPFRSMVRV
APQVALQSYRSYVSKQVYQVYDEHLRQAFSFLVGNPNPATSIIYALIHALLERW
GWVFRGDTGALINKGLQVLFKDLGVEVTLTSVSRITETANGRVSAAVVAEDRRPAAD
IVASNAVHTYTRDLKDEPLARPTAQLSRKRFSMSFLVYVFLRRRHPHKLHII
FGRRYRELINIEFKGPALEDESILYHAPSVDPSLAQGCSTYTVLSPVPHLAAPRI
DWSVGPRYRDLIYLEARILPGLKSLATCRIFTPQDFNTELNHAGSAPSLPIL
TQSAYFRAHNADKIKGLVYAGTHPGAGIPGVVGSAAKATARVILEQRELAQA"
4144. .5151
/gene="crtB"
CDS
4144. .5151
/gene="crtB"
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="AAF78202.1"
/db_xref="GI:8650418"
/translation="WTLRDNHDLARLSEAVINGSKSPAASKLFDSTRASVHLLYA
WCRHDDVIDGDLGIRGVGAPGQIGLQMLRDOTAQLGAPMPDPVFGQLQRVV
QBAHAPHHVFEELGFMVDVREYETLSETLDYCYHAGVGVYVMSAINGAREEAT
LDRAADGLIALQLTNIARDVIEDAQTRMPLQWLCEAGVPAEVAISPOHQAQVFRV
VARLDVAEQFEASEAGIARLPVRCAMAVETARVVYQIGREVMKRGPGAWDARIAT
TGAQKLGAIQRSALTIGLTRTWKVPAREHNLWTRPKARRAASDAQSPANACSNRQ
Y"
5123. .5899
/gene="crtW"
CDS
5123. .5899
/gene="crtW"
/codon_start=1
/transl_table=11
/product="beta-carotene ketolase"
/protein_id="AAF78203.1"
/db_xref="GI:8650419"
/translation="MHAATAKATEFGASRRDDARORRVGLTLAAVILAAWLVLHVGLM
FEMPLTLHSLIPALPLVLTWLYVGLFTIAHDMHSGSLVLPKQVNRRIQGLCLFLY
AGPSFDALNEHKKHRRPGTADDFDEVPFHGFHWFFSFFLHYFGWKQVALLIAV
SLVYQLVFAVPLQNLILLFWALPGLLSALQLFTFGYLPKHPKATQPFADRHNRATSEFP
AWLSLLTCFHFHFHHEHLHPDAPWRLPFIKRRALERRD"

```

ORIGIN

Query Match 15.8%; Score 1105.8; DB 1; Length 6635;

Best Local Similarity 58.4%; Pred. No. 5.6e-222;

Matches 2082; Conservative 0; Mismatches 1417; Indels 66; Gaps 6;

QY	2502	GTGCTGGCGGAGCGACCTGTATGCGCACGCAATACATGTGATTTTGGTCGGTGTGGAC	2561
DB	1426	GAGCGGCTGCGTTGTGTCATGATCGAGATGCCAGCTCATGTCATCGCGCGGCTC	1485
QY	2562	TGGCAATGCGTTGATTCGCGTTCGCGTCAATTGCGAGCCACAACACTGAAATCCCTGT	2621
DB	1486	TTGCGCGCTGCGTATGCGCTTTCGCGCTCACCGACGCGCGGATCTGCGCGTCGTCA	1545
QY	2622	TGCTGAGAGCGATCGGCATCGGCAGCAATCATACCTGGTCGTTTCAATCACAGCGATC	2681
DB	1546	TCATCGAAGGCTCGGCCAGCATCGCGCGCAATCACACCTGGAGCTTCTTTGGAACCGATA	1605
QY	2682	TCAGCGCCGAAACAACCTTCGCTGGCTGCAACCGCTGATTACCGTGGTGGTCAAGTTATC	2741
DB	1606	TCTCGTCCGACCGACGCGCTGGCTCGAGCGGCTCGTCGGTCAATCGCTGCCCGGTTATG	1665
QY	2742	AGTTCGGTTCCTCGCTGCGCCCAATCTCGAGCGGGATTTATGTTCCATCGCATCAG	2801
DB	1666	AGTTCGGTTCGCCGAAATGCCATCCGTCCTGACCGCCCTACCTCTCTAGAGCTCGA	1725
QY	2802	GCGATTTTGGCCGCCATCTTTACGCGGCGATGGGTGACGATCTGTGGACAAACACAGCGG	2861
DB	1726	CAGCGCTGCTGCGCGAGGTCGAGCAGCGTTTCCGAGAGGATCTCGCGCAGCGACGG	1785
QY	2862	TACAACAGGTAAACCCAGCAGGTGACGCTGGCGCGATGGCCGTGAACCTTGTGTGCGCAAG	2921
DB	1786	CCATCTCCGCAACGCGACAGACCATGTCGTCTGGAGGGCGCGCACCTTTCGCGCGGCCCT	1845
QY	2922	TGGTGAATGATGTCGCGCCCTGCGAGCGCGACGCGACACATCTGAGCTGGTGTATCAGGTGT	2981
DB	1846	GCGTGTGATCGATCGCGCGCGCGCGCGCGCGGTTCGCTCTCGGGTTTCAGAACT	1905
QY	2982	TTCTTGGACAAGAGTGGCAGCTGGCGCAGCGCGACGCGCTGCGAGCGCGATCTCTGATGG	3041
DB	1906	TCCTCGGCTCGAGGTGGGTGGCGCGCGCGCGCGCGCTCGATGTGCCATCTGATGG	1965
QY	3042	ATGCCACCGTTCGATCAGCAAGCGGTTATCGTTTGTCTACACGCTGCGCGCTCAGCGCGG	3101
DB	1966	ACGCGACGCTCGCGCAGAGCGAGCGCTATCGTCTGCTTACACGCTGCGCGCTCGACCCCTC	2025
QY	3102	ATCGGCTATTGATGAAGATACCCATAGGTAAACAGCGCGCGCTGGCGGAGAACACCG	3161
DB	2026	AGCGGCTGTTGATCGAGGACACCTACTATAGCGACGCGCGGAGTTGCCCGAGCAGGTGC	2085
QY	3162	CTCGTCAGCATCGCCGACTATGCCAATCAGCAAGGCTGGACGCTGAGTACGCTGCTGCTC	3221
DB	2086	TGCTACGCGATCGCGGCTACGCGCTCGCCAGCGCTGGCAGATCGCCGAGATCATCC	2145
QY	3222	GTGAAGACCGGCATATTACCGATTACCTTGAGCGGCAACATCGATCGATTCTGSCAAC	3281
DB	2146	CGCGGAGCAGGCTGCTGCTGCGCGCTCATCTTGGCGGCGATCCGCTCGGCGTGGTCAGCA	2205
QY	3282	AGCAGCGCGGCAAGCTGCGAGCGCTGCGCGCGCGGCTGTTTATGCCACACCGGTT	3341
DB	2206	AGCCGACAGCGCGCGCGCGTGGGTCTCGCGCGCTTCTCGTGCATCCGACGACCGGCT	2265
QY	3342	ACTCTTCGCGTCCGCTGCGCTGCGCTAGCGGAGTTGCTAGCAGCGCTG-TTGCCACCGAT	3400
DB	2266	ATTCTCTCCGACCGGTCGCGTCCGCTGCTGACGCGCGGCTTGGCAGACGGG	2325
QY	3401	GCCCTCAGCGCTCA-----GCCAACATATCGAACGCTTTTGCCTTCAGCAGTGGC	3449
DB	2326	GCGCGCTCTCGAGCGCTGACGCGCGGAGACGATCGATGGCTACGCGCGGACGATCTGC	2385
QY	3450	GCGAACACGGAATTTTCGCTGCTGCTAAACCGCATGCTGTTTGGCGCGGTAAAGCGCAGC	3509
DB	2386	GCGCGCGGCGCTACTATCGCTTCTTGAACCGGATGCTTCAAGCGCGCGCGACGCTCGC	2445
QY	3510	AGCGTGGCGGCTGATGCAACGTTTTCACGCGCTGCGCGGTTAATACCGCTTTT	3569
DB	2446	AGCGTACCGGAATCTCGACAGATTTTACGGTCTCGATCAGGCCCTGATCGAGCGGCTCT	2505
QY	3570	ACGCGCGGCAACTGCGGCTCGCGGATAAAACGCGGAT-----TCTGTGCGGCAAGC	3620
DB	2506	ACGCGCGCGGATCCAGCGCGGAGACAGCTGGCGGCTCTCATGCATATGCTGATGAAGC	2565
QY	3621	GCGCGGTGCCCATCGGTGAACGCT-----GCGCGCGTGTGTAAT	3662
DB	2566	GCGCGATCCGATCTCGTCCGCGCTCGCTGCTGCTGCTGAGCGGAGCGCTTCAAGACCC	2625
QY	3663	CTGTGCAACACGAGGAGAAAAATGAACGACATTATGTTGTTGGCGGCGGTTTGGCG	3722
DB	2626	CATGAGTCAAGAGATGCAAGAACGCGCAACAGCAGTGTGTGATCGGTTTGGGTGG	2685

QY 3723 CTTGCGGCTGGCGATTCGCTGCAAGCGCGGGGCATACCAACACCTTACTCGACAGCG 3782
Db 2686 GCTTTTCGCTCGCATACGAGCTTCAATCGCGGGATCGCAACACGCTCGTGGAGAGCG 2745
QY 3793 CGACAAACGGGGCGAGCGGCTATGTGTTGAGGACAGTGGCTTTACTTCGATCGCGG 3842
Db 2746 GCACAAGTTTCGGCGCGCGCGCTATGTCTACAGAGGAGGACGCGCTTCACTTCGACGCGG 2805
QY 3843 ACCCACGGTGATCACCGATCCAGCGCCATCGAAGAGTTGTTCAAGCTGGCGAGGAAATC 3902
Db 2806 GCCCACTGTATACCGATCGCACTGCTGCGAGAGCTGTTTCGCGCTCTCGGGCGCAA 2865
QY 3903 GCTCAGCGATTACGTGAGCTGATGCGGTAAACGCGCTTCTATCGCTGTGCTGGGAAGA 3962
Db 2866 GCTCAGAACTATGTGAGCTGATGCGGTCTCCCATTTCTATCAGCTGCGTGGGAGGA 2925
QY 3963 TGGCAAAACAGTTGATTAGCAAAATANTACGCGCGTGTGAGGACAGAGATCGCCACGTT 4022
Db 2926 CGCGGCCCACTTTCGACTAGCTCAACGACCGAGCGGAGCTGGAGCGCCAGATCGCGCGGTT 2985
QY 4023 CAATCCGCAAGATAGAGGCTATCGTCAATTTCTTGCTTATTCACGTGAAGTATTTAG 4082
Db 2986 CTGTCCGGCTGATGAGCGGCTATCGGGCTTCCGTCGTATAGGAGCGGCTCTGGA 3045
QY 4083 AGAGGTTTATCTGAACCTCGGCAACGCTGCGGTTTCTGCAAGTGTGCTGATGCTCGCGT 4142
Db 3046 GGAGGGCTACGTCAAGCTCGGCCACGTCGCGTTCCCGGATTTCCGGAGCATGGTTCGTGT 3105
QY 4143 CGCGCGCAGTTGGAGCTCTGCAAGCATGGCGAGCTCTACAGCATGGTGGCGAAAT 4202
Db 3106 GCGCGCGCAGCTGGTGGCGCTGACAGCTATCGCAGCGTCTATGCAAGGTTGTGCAATA 3165
QY 4203 TATTGAGGACGATCATCTGCTGAGCGGTTTCTTCCACTCATTTGCTGGTGGCGGTAA 4262
Db 3166 CGTCTCCGACGAGCATCTGCGCAGCGCTTCACTTCCACTGCTGCTGCTGCGCGCAA 3225
QY 4263 TCTTTTGCACGTCATCGATCTATACCTTAATCATGCTGAGCGTGAATGGGCGT 4322
Db 3226 TCCGTTTCGCGACGTCCTCGATCTACGCTGATCCATCGCTGAGCGCGCTGGGAGT 3285
QY 4323 GTGGTTTCGCGCGCGGACCGGCGCGTGT---GCAGGCAATGGCGCGACTGTTGA 4379
Db 3286 TCGGTTTCGCGCGCGGACCGGCGCGCTCATCAACAAAGGGGCTCGTCCAGCTGTTCAA 3345
QY 4380 GACATTGGCGCGGAGCTGTTACTGAATGCCAAGTGAGCCAGCTGGAACACGAGCGCAA 4439
Db 3346 GGATCTCGCGCGGAGGTGACGCTGTGACAGCTGAGCGCGATCGAGCGGCAACGG 3405
QY 4440 TCGCATTAGCGCGTTTCACTTAAAGGCGGAGCGACGCTTTCGATGCGCGCGCTGCGCTC 4499
Db 3406 GCGGTCAGTGCCTGCTGCGCGAGGAGCGCGCTCGGTTGCGCGCGACATCGTGCACG 3465
QY 4500 CAATCGCGAGTGTGATACCTACGACAACTGCTTCCGCAACCATCGCTGGCAATGNA 4559
Db 3466 CAACCGCGACGCTGTTACTACCTACCGGATCTCTGAAGGACGAGCGCTGGCGCGCC 3525
QY 4560 AGTGGCATCGCTGAAGCGTAAGCGATGAGCAACTCGCTGTTGTTGCTACTTATTTGG 4619
Db 3526 GACGCGGCATCTTTGATGCGCAAGGCTTCAAGATGTCGCTGTTGCTGATCTTTGG 3585
QY 4620 CTTGAATCAGCGCGATGAACTGCTGCGCAACACCGCTGCTGTTTGGCGCGCGTTATCG 4679
Db 3586 CTTGCGCGCGAGCATCGGAGCTCAAGCACCACATCATTTCTGTTCCGCGCGCTACCG 3645
QY 4680 TGAGTTGATGATGAGATTTTCAACAGCAGCAGCTGGCAGAGATTTTCACTTTACT 4739
Db 3646 CGAGCTGATCAACAGAGATCTTCAAGGGGCGCGCTGCTGAGGATTTCTCGCTGATCT 3705
QY 4740 GCACGCGCGCTCGACGAGCATCTCGTGGCACCGCGCTGCGGAGCTTTTATGT 4799
Db 3706 GCACGCGCGAGGTTGACCGATCCGTGCTCGCGCCACAAAGGCTCGACGACCTATTACGT 3765
QY 4800 GTTAGCGCGGTCGCGCATCTCGGCACCGCTGACATCGACTGGCAACAGGAAGGACCGG 4859

Db 3766 GCTCTCGCGGTCGCGCATCTCGCGCGCACCGATCGACTCGAGCGTGGAGGCGCGCG 3825
QY 4860 CTTGGCGGATCGAATTTTTCCTTATCTGAGACGACATACATGCGGGATTAAGTCACA 4919
Db 3826 CTATCGGACCGCATCTCGACTATCTCGAGGCGCGCATCTGCGGGGCTGAAGTCGGA 3885
QY 4920 ATTAGTGACACACAGAAATGTTTACCGCTTTTTCGCGACACGCTGCAATGCCATCA 4979
Db 3886 CCTGCCACCTGCGCATCTTCAAGCGCAGGATTTCAACACCGAGCTGAATGCGCATCT 3945
QY 4980 CGGCTCGCGGTTTTCGCTGAGCGGATTTTGAAGCAAGAGCGCTGGTTCGCGCGCATAA 5039
Db 3946 CGGCTCGCGGTTTTCGCTGAGCGGATCTTCAAGCAGAGCGGATTTTTCGCGCGCCACA 4005
QY 5040 CCGGATCGCGCATATCAGCAATCTCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 5099
Db 4006 CGCGACACAGAGATCAAGGGGCTGTATCTCTGTCGCGCGCGGACCCATCTCGGGCGCG 4065
QY 5100 CTTGCGCGGCTGATCGGTTTCGCGCAAGCGCACCGCGCTGATGCTGGAGGATTCGCGC 5159
Db 4066 CATTCGCGGCTGCTGCGCTCGGCCAAGCGCACCGCGGAGTCTCTCGAGGACGAGCG 4125
QY 5160 GCAATGAATCGACAGCTTTTACTTGAAGCAAGTAAAGCAACCAATGGCGGCTG----- 5210
Db 4126 CGAACTCGCGGAGGCGCAATGACGTGCGCGACCAATGACCATCTGCGCGCGCTGAGCGAA 4185
QY 5211 -----GGCTCGAAGAGTTTTCGCCACCGCGCAAGCTGTTTGAATGACCG 5255
Db 4186 GCGCTATCCGCAACGCTCGAAGAGTTTCGAGCGCATCGAAGCTGTTTGAATGACCG 4245
QY 5256 ACGCGCGCGACGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5315
Db 4246 ACCGTGCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4305
QY 5316 GGGCAACGCTGGGCGAAGCGCGCACGAGCATGCGCTGGAAGCGCGACGCGAGCTATG 5375
Db 4306 GGAAGGATCTCGAAATTCGCGAGGCGCTGCGCGCGCTGCGCGCGAGATCGGAGCTTTG 4365
QY 5376 CAGCATCTGCAAAATGAAACCGCGCTACAGCGCGCGCATCGATGAAACCGCGCG 5435
Db 4366 CAGATCTGCGCGACAGACGCGCGAGCGCTGGAGGGGCGCGCATGCGCGATTCGCTG 4425
QY 5436 TTTAGGCGGTTTTCAGGAGTGGCGATCATTCACAGCTGCGCGCAACAACTGCGCGTTGAT 5495
Db 4426 TTTCAAGGATTTGACGCGCTGCTGAGGACGACGCGATTTCCGCAACATCACTGTTGAG 4485
QY 5496 CATCTGGAAGCTTTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5555
Db 4486 CTGCTCGACGCTTCGCCATGATGTGACGCGCGCGAATACGAGACGCTGAGCGAGACG 4545
QY 5556 CTGCTTACTGCTATCAGCTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5615
Db 4546 CTGACTACTGCTATCAGCTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4605
QY 5616 GTGCGCGACGAGCGGCTGCTGATCAGCTGCGATTTAGGAGCTGGCGTTTCAAGTCACT 5675
Db 4606 GCTCGAGGAGGCGAGCTGGAACGCGCGCGCGATCTCGCATCGCGCTGAGCTTACC 4665
QY 5676 AACATTGCGCGGACATTTGTAGAGATGCGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5735
Db 4666 AACATCGCGCTGATGAGGATGCGGAGTCCAGACCGCGCGCATGTATCTGCGCGCAAA 4725
QY 5736 TGGCTCGATCAGGCGGATTTACGCGCGATACGCTGACTGACCGCAACCACTGTCAGCG 5795
Db 4726 TGGCTGCGAGGCGCGGCTGCGCGCGCGCGGCTGCGGAGTCCGGAACCGAGCATCGCGAGG 4785
QY 5796 CTCGCTCAGCTGGCAGCGGCTTTAGTGGGAGGCGGAAACCTTATATCACTTCGCGCGGA 5855
Db 4786 GTGTTCCGCTGCTGCGCGGCTGCTGATGTCGCGGAGGAGTTTTACGAGGCGAGCGAG 4845
QY 5856 TCGGTTTACCGGTTTACGCTGCGCTGCGCTGCGGCTGCGCTGCGCTGCGCTGCGGCTGCT 5915

Db 4846 GCAGGATCGCCCGCTCGGTGCGTTCGGCTGGCGGTGGAGACGCCGCGCTCGTC 4905

QY 5916 TATCGCGAAATGGCGTCAAGTTCAGCAGCGCGGTGTCACAGCTGGGATTCACGGCAG 5975

Db 4906 TATCGCGAGATCGCGCGCGAGTGATGAAGCGCGCTCCGCGCGCTGGAGCGCCGCATC 4965

QY 5976 CCACACAGTAAGGTGAATACTGG 6000

Db 4966 GCCACGACAGCGCGCAGAGCTCG 4990

RESULT 12

LOCUS PNCRTWZYI 6335 bp DNA linear BCT 15-SEP-1999

DEFINITION Paracoccus marcusii crtW, crtZ, crtY, crtI, crtB & crtE genes.

ACCESSION Y15112

VERSION Y15112.1 GI:5912291

KEYWORDS astaxanthin biosynthesis; beta-carotene C-4 oxygenase; carotene hydroxylase; crtB gene; crtE gene; crtI gene; crtW gene; crtY gene; crtZ gene; geranylgeranyl pyrophosphate synthase; ketolase; lycopene beta-cyclase; phytoene desaturase; phytoene synthase.

SOURCE Paracoccus marcusii

ORGANISM Paracoccus marcusii

Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales; Rhodobacteraceae; Paracoccus.

REFERENCE 1

AUTHORS Harker, M. and Hirschberg, J.

TITLE Carotenoid biosynthesis genes in the bacterium Paracoccus marcusii MH1

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 6335)

AUTHORS Hirschberg, J.

TITLE Direct Submission

JOURNAL Submitted (12-SEP-1997) J. Hirschberg, The Hebrew Univ of Jerusalem, Dept of Genetics, Life Sciences Institute, Jerusalem 91904, ISRAEL

FEATURES

source

1..6335

/organism="Paracoccus marcusii"

/mol_type="genomic DNA"

/strain="MH1"

/db_xref="taxon:59779"

409..1137

/gene="crtW"

409..1137

/gene="crtW"

/function="astaxanthin biosynthesis"

/codon_start=1

/transl_table=11

/product="beta-carotene C-4-oxygenase (ketolase)"

/protein_id="CAB56059.1"

/db_xref="GI:5912292"

/db_xref="GOA:Q9RLH7"

/db_xref="TrEMBL:Q9RLH7"

/translation="MSAALPKADLTATSLIVSGGIIAAILAHVHALWFLDAAHPHILAVANFLGTLWLSGLFTIAHDAMGVSVPGRANAAWQQLVLYAGFWRKMLVKVNAHRRAGTDDPDHFGGVPVRYARTIGTYFGWRGGLLPVITVVALILGDRWVYVFLPLSILASIQLFVGTGTLPHRPGHDAFDPRNARSRSISDPVSLTLCFHFGGYHHEHLHPVTVPWRLPSTRTKGDTA"

1134..1622

/gene="crtZ"

1134..1622

/gene="crtZ"

/function="astaxanthin biosynthesis"

/codon_start=1

/transl_table=11

/product="carotene hydroxylase"

/protein_id="CAB56060.1"

/db_xref="GI:5912293"

/db_xref="GOA:Q9RLH6"

/db_xref="TrEMBL:Q9RLH6"

/translation="WTNFIIVATVLVMTATYSVHRWIMHGPLCGWGHKSHHSHDHALEKNDLGLVFAVTATVLTFTVGNIWAPVLWIALGMTVYGLIYFLHDGVVHORWPP

RT"

RYIPKGYARELYQAHRLHHAVERGDHCVSFGFIYAPVPPDKLQDKLTSGVLRABAEQ

1619..2779

/gene="crtY"

1619..2779

/gene="crtY"

/function="astaxanthin biosynthesis"

/codon_start=1

/transl_table=11

/product="lycopene beta-cyclase"

/protein_id="CAB56061.1"

/db_xref="GI:5912294"

/db_xref="TrEMBL:Q9RLH5"

/translation="MTHDVLLAGAGLIALLAALRAAPDLRVLLDHAAGPSEHGTWSCDDPDLSPHLKRLPRLANMPDQVRFPRHARRLATGSGLDGALADAVVRSGAEIWDSDIALLDAQATLSCGTREIAGSVLDGRGPHSRHLTLGFHKFLGVEIETDRPHGVPRLVMDGTVTQDGYGFIYLLPFSRTRILTEDTRYSDGGDLDDALAAASQDYAROCQTCAEVRERRGILPIALAHDAAGFADHAEQPVVGLRAGFPHFVTCYSLPYAQVADVAGLSGPPCTDALRCARIDYADRARORFLRLNMLRFGCAGPDRRYTLQRLYMPHGLIERFYAGRLSVADQLRIVTGKPIPLGTAIRCLPERPLKENA"

2776..4281

/gene="crtI"

2776..4281

/gene="crtI"

/function="astaxanthin biosynthesis"

/codon_start=1

/transl_table=11

/product="phytoene desaturase"

/protein_id="CAB56062.1"

/db_xref="GI:5912295"

/db_xref="GOA:Q9RLH4"

/db_xref="TrEMBL:Q9RLH4"

/translation="MNAHSPAAKTAIVTIGAGFGGLALAIRLQSGAIATTLVEARDKPGGRAYVMDHQHVFDAGPTVITDPALKELWALTQDMARDVTLMVSPFYRLMWPGGKVFVYNEADQLERQIAQFNPDLEGRYFRDYAEVYQGVYKLTGTPFLKQMLKAAPALMKLAYNSHAKVATFKDPYLROAFSVHTLLVGNPFSTISIVALIHALLERRGCVWFAGKGTNLQVAGMVALFERPGGOMLNAAKVARLETGEGARTTCVTIADGSLRADMVASNGDVNHTRDLLGHTARGQSRAKSLDRKWSLFLVLFHGLREAPKDIATHTILFGPRYKELVNEIFKPKLAEDFSLYXHSPTTDDPNAPGMSHTHYVLAVPPHUSRAEIDWAGGPRYADRILAFLEERLIPNLRLNLTTRITFTPADFASELNHHGSAFSEVPILTQSAWFRPHNRDKTIRNFYLVGAGTHPGAGIPGVVGSKATAQVMLSDLASA"

4278..5192

/gene="crtB"

4278..5192

/gene="crtB"

/function="astaxanthin biosynthesis"

/codon_start=1

/transl_table=11

/product="phytoene synthase"

/protein_id="CAB56063.1"

/db_xref="GI:5912296"

/db_xref="GOA:Q9RLH3"

/db_xref="TrEMBL:Q9RLH3"

/translation="MSDILVLTSTAITQGSOFATAAKLMPGPGRDITVMLYAWCRHADDVIGQALGSRPEAVNDQARLDGLRADTLAALQDQGVPTPPAALRAVARRHDPFQAWPMDLISGPAIDVEARYTRILDVLEYSYHVAGILGVNMAVWGVRDHPVLDRACDLGLAQTLTHADVIDYARIGRCYLPDMDWDHAGARVDGVPSPSELYTIVRLDRAEPPYPSARGLADLPKPCAWSIPAAAFIYRAIGLPIRKGGPEAYRQRISTSKAAKIGLLGICGMDVADHACGSGVSRQDLWTRPPHA"

complement(5189..6064)

/gene="crtB"

complement(5189..6064)

/gene="crtB"

/function="astaxanthin biosynthesis"

/codon_start=1

/transl_table=11

/product="geranylgeranyl pyrophosphate synthase"

/protein_id="CAB56064.1"

/db_xref="GI:5912297"

/db_xref="GOA:Q9RLH2"

/db_xref="TrEMBL:Q9RLH2"

/translation="MRRDVNPHITLLOTRITEEIAQGGFAGVLIQPLGAAMAALSSGKRF

RGMLMLLAAESGGVCDTIVDAACAVENVHAASLIFDDLPCMDDAGLRGQPATHVAH

GESRAVLGGIALITEAMALLAGARGASVTRVLAQLVRLISRLSPQGLCAGODLHLHAA
KNGAGVEQDLKTVGLPIAGLEMLAVIKEDAEQOTQMDIFGROLGRVPSYDLDLE
VYGOAALGKDTGDAAPGPRGLLAVSDLKVSRQYEASRAQLGHNVSRLQAPB
TEALLERVLVPAARA"

ORIGIN

Query Match 14.9%; Score 1041.2; DB 1; Length 6335;
Best Local Similarity 57.4%; Pred. No. 2.3e-208;
Matches 2059; Conservative 1; Mismatches 1479; Indels 51; Gaps 8;

Qy	2456	GGTGTCCGCGCAGCATCGTGGAGCAGCGCTGTGCCAGCAGCAAGTCGTGTGGGGAGGC	2515
Db	1547	GGCCCGGTTCGACAGCTGAAGCAGCACTGAAGCAGCTCGGCGGTACTGTGGGGCGCAGGC	1606
Qy	2516	GACCTGATGCGCACGCAATACGATGATTTGGTTCGGTGTGGACTGGCGAATGGCTTG	2575
Db	1607	GCAGAGCGCAGCTGACCCATGACGTGTGTGGCAGGGGGGGCTTTGCCAAGCGGGCTG	1666
Qy	2576	ATTGCGCTGCGTCTCGGTCAATTTGAGCCACAACTGAAATGCTTGTCTGGAGAGCGAT	2635
Db	1667	ATCGCCCTGGCGCTGGCGCGCGGCCGACCTGCGGTGCTGCTGTGACCATCGG	1726
Qy	2636	GGCGATTCGGCAGGCAATATACCTGGTGTGTTTCAACAGCGATCTCAGCGCCGAAACA	2695
Db	1727	GGCGACCGTCAGAAAGGCCATACCTGGTCAATGCCACGACCCCGATCTGTGCGCGCACTGG	1786
Qy	2696	CTTCTGCTGGCTCAACCGCTGATTACCGTGTGGTGTGCTGAGTTATCAGTGTGTTTCT	2755
Db	1787	GTGGCGCGCTGAAGCCCTGCGCGCGCAACTGGCCGACCGAGAGGTGCGTTCCTCC	1846
Qy	2756	GCGCTGCGCGCAATCTGACGCGGATTTATTTGTCATCGCATCAGGCGATTTTCCCGCG	2815
Db	1847	CGCCATGCCCGCGGTGSCCACCAGGTACGGGTGCTGGAGCGGCGCGTGGC---G	1903
Qy	2816	CATCTTTACCGCGGATGGGTGACGATCTGTGGACAACAACAGCCGTACACAGTAATA	2875
Db	1904	GATGCGGTGTGCGGTGCGGCGCCGAGATCCGCTGGGACAGCAATCGCTTGTGAT	1963
Qy	2876	CCACAGGTGACGTGCGCGATGGCGTGAATCTGTGCGCAAGTGTGATTCATGAT	2935
Db	1964	GGCAGGGGGACACTGTCTGCGGCACCGGATCGAGCCGGCTCGTCTGGAGCGG	2023
Qy	2936	CGCGCTGACGCGCAGCCACATCTGACGCTGGGTATCAGGTGTTTCTTGGACAAG	2995
Db	2024	CGGGGCGCCACCGCTCCCGCATCTGACCTTGGGTTTCCACAAATTCCTGGGTGTGAG	2083
Qy	2996	TGGCAGCTGGCAGCGCGCCGCTGAGAGCGGATCTGATGGATGTCACCGTCGAT	3055
Db	2084	ATCGAGACGACCGCCCCCAGCGCGTGGCCCGCGGTGATCATGACGGGACAGTCACC	2143
Qy	3056	CAGCAGCGGGTTATCGTTTGTCTACAGCTGCGCGCTCAGCGCGGATCGGCTATTGAT	3115
Db	2144	CAGCGGACGGGTACGGCTTCATTATCTGCTGCCCTTCTCTCGGACCGCATCTGATC	2203
Qy	3116	GAAGATACCATTAACGATACGACCGCGCTGGCGGAGAAACACCGCTGTGACGACATC	3175
Db	2204	GAGGACACGCTATTCCGATGGCGCGATCTGGAACGACGCGCTGGCGCGCGTCC	2263
Qy	3176	GCGGATATGCCAATCAGCAAGCTGAGCTGAGTACGCTGTGTGTAAGAGCAGCGC	3235
Db	2264	CAAGATCTATGCCCGCAGCAGGGTTGGACCGGGCGGAGTTCGGGCCGAA---CGGGC	2320
Qy	3236	ATATTACGATTAACCTGAGCGGCAACATCGATTCGATTTCTGG---CAACAGCAGCGCGC	3292
Db	2321	ATCTGCTATTGCGTTGGCCCATGACGGCGGGCTTCTGGCCGATCACCGGAGGG	2380
Qy	3293	CAAGCGTGACGGCTGTGGCGCGGGCTGTTTATGCGCACACCGGTTACTTCTTCCG	3352
Db	2381	CTGTGTTCCGTGGGACTGTGCGCGGGGTTCTTTCACCCGGTCAACCGGCTATTGCTGCC	2440
Qy	3353	TCCGCGTGTGCGCTAGCGGATTTGGTAGAGCGCTGT---TGCCACCGATGCCCTCACG	3409
Db	2441	TATCGCGCGAGGTGGCGGACGTGGTGGCGGGCTGTCCGGGCGCGCCGCGCACCGACG	2500

Qy	3410	CTCAGCCAAACATATCGAAACGCTTTGCGCGTCAGCAGTGGCGGAAACAGACGATTTTCCGT	3469
Db	2501	CTGGCGCGCCATCCCGGATACCGATCGACGGGCAACCGCGGACCGCTTCTTCCGC	2560
Qy	3470	CTGTAAACCGCATGTGTTTTTGGCGGTAAGCCGACAGCGCTGGCGGTGATGCA	3529
Db	2561	CTGTAAACCGCATGTGTTTTCCGCGCTGGCGGCCCGACCGGCTATATACCTGCTGAG	2620
Qy	3530	CGTTTTTACCGGCTCGATCGCGGTAAATAGCGCTTTTACGCCGGGCAACTGCGCTG	3589
Db	2621	CGGTATACCGCATGCGCGATGACTGATCGAACGGTTCATGCCCGCGGCTGAGCGTG	2680
Qy	3590	CGCATAAAAACGGGATTTCTGTGGCAACGCGCGGTGCCCATTCGTGTGAAGCGCTCGC	3649
Db	2681	CGCATCAGCTGCGCATCGTGACCGGCAAGCTCCCATTCCTTTGGCACGGCCATACGC	2740
Qy	3650	GCGTGTGTGAATTTCTGTGGAACGAGGAAAGAAAATG-----A	3688
Db	2741	TGCTGCCCGAAACGCTCCCTGCTGAGGAAACGCAATGAACGCGCATTCGCCCCCGGCCA	2800
Qy	3689	AACGCACTTATGTATGGCGCAGGCTTTGGCGCTGGCGCTGGCGATTCGCTGCAAG	3748
Db	2801	AAACCGCATCTGATCGCGCAGGCTTTGGCGGCTGGCGCTGGCGCATCGCGCTGAGT	2860
Qy	3749	CGGCGGCATACCAACCACTTACTCGAGCAGCGCGCAAAACCGGGCGGACGCGCTATG	3808
Db	2861	CGGCGGCATCGCCACCACTTGTGAGGCGCGCAAGCCCGCGGCGCGCTATG	2920
Qy	3809	TGTTTGAAGCAGTGGCTTTTACCTTCGATCGCGACCCACGCTGATCACCGATCCACGC	3868
Db	2921	TCGCGCATGTCAGGCGCATGCTTTTCAGCGCGGCGCGACTGTCTATCACCGACCCGATG	2980
Qy	3869	CCATCGAAGAGTGTCTCAACGCTGCGAGGAAATCGCTCAGCGATTAAGCTCAGCTGATGC	3928
Db	2981	CGCTCAAGGAGCTGTGGCGCTGACCGGCGAGGACATGGCGCGGACGCTGACGCTGATGC	3040
Qy	3929	CGGTAAACCGCTTCTATGCGCTGTGTGGGAAGATGGCAACACAGCTTGTATACACATA	3988
Db	3041	CGGTGTGCCCTTCTATCGGCTGATGTGGCGCGGCGGAGGTTCTTGATTTACGTGAACG	3100
Qy	3989	ATCAGCGCTGCTGAGCAGCAGATCGCCAGTTCAATTCGCAAGATGTGAAGGCTATC	4048
Db	3101	AGGCGCATGACGTGGAACGCCAATTCGCCAGTTCAACCCGAGACGCTGGAAGATACC	3160
Qy	4049	GTCAATTTCTTCCCTATTACGTGAAGTATTTAGAGGGTTATCTGAAATCTCGCACGG	4108
Db	3161	GTGCTTCCGTGATTAAGCTGAGGAGGTGTACAGGAGGGCTAGCTCAAGCTGGGACCG	3220
Qy	4109	TGCGGTTTCTGAGGTGCGTACATGTGCGGCTGGCGCGCGAGTTGGAGCTGTGCAAG	4168
Db	3221	TGCGCTTCTCAAGCTGGGCGCAATGTCTCAAGGCGCGCGCGCTGATGAGCTTGAGG	3280
Qy	4169	CATGCGCAGCGCTTACAGCATGTGGCGAAATTTATTCAGGACGATCATCTGCGTCAAG	4228
Db	3281	CCTACATTCGTCCATCGCAAGGTTCGCACTTCATCAAGAACCCCTATCTGCGGACAG	3340
Qy	4229	CGTTTCTTCCACTCATTTGCTGTGGCGGTAAATCTTTTGGCAACGTATCGATCTATA	4288
Db	3341	CGTTTTTGTATCACACGCTGTGTGGCGGGAATCCCTTCTCGACACGCTGATCTATG	3400
Qy	4289	CGTTAATTCATGCGCTGAGCGTGAATGGGGCTGTGGTTTCCGCGCGCGCACCGCGC	4348
Db	3401	CGCTGATCCAGCGCTGAGCGCGCGCGGGGTCTGGTTCCGAGGCGCGCACCAACC	3460
Qy	4349	CGCTGCTGAGCGCATGGCGGACTTGTTCGAGGACTTGGCGCGCGAGCTGTTACTGAATG	4408
Db	3461	AGCTGCTCGGGCATGTTCGCGCTGTTCGAACGCGCTGGCGCGCAGATGATGCTGAACG	3520
Qy	4409	CCGAAGTGAAGCGCTGGAACACGCGGCAATCGCATTTAGCGGCTTCACTTAGAGGGCG	4468
Db	3521	CCAGGTGCGCGGATCGAGACCGGAGGCGCGGACCAACCGGCGCTCACCTTGGCGGACG	3580

/note="Agrobacterium aurantiacum is distributed from
Marine Biotechnology Institute as Paracoccus sp. MBIC1143.
Its 16S rRNA sequence is shown in accession
AB008114.-cloning vector: pAK32"
139. .867
/gene="crtW"
139. .867
/gene="crtW"
/citation={2}
/codon_start=1
/transl_table=11
/product="beta-Carotene ketolase"
/protein_id="BAA09591.1"
/db_xref="GI:1136631"
/translation="MSAHLPAKADLTATSLVSGGIIAALVHFWLWFLDAAAHPI
LAIAHFLGLTWSVLGFLTIAHDMHGSVVRPRANAMQGLVWLWVAGFSWRKQIVK
HMAHHRAGTDDDDPFDHGGVVRVAREIFTYFGWREGLLLPVITVVALTLGRDWMY
VVFPLPSILASIQLFVFTGWLPHRPGHDAFPDRHNRSSRISDPVSLITCFHFGGYH
HEHHLHPTVPMWRFLPSTRTKGDTA"
864. .1352
/gene="crtZ"
864. .1352
/gene="crtZ"
/citation={2}
/codon_start=1
/transl_table=11
/product="beta-Carotene hydroxylase"
/protein_id="BAA09592.1"
/db_xref="GI:1136632"
/translation="MTNFIIVVATLVNMLTAYSVHRWIMHGPGLGWKHSHEHHDH
ALEKNDLVGLPAVATLFTVGTWAPVLMIALGMTVYGLIYFVLHDGLVHQRWPF
RYIPRKYARRLYQAHLHHAVEGRDHCVSFGFIYAPVPVDKLLQDLKMSGLVRAEAO
E"

gene
CDS

gene
CDS

gene
CDS

gene
CDS

gene
CDS

1349. .2509
/gene="crtY"
1349. .2509
/gene="crtY"
/citation={2}
/codon_start=1
/transl_table=11
/product="lycopenase cyclase"
/protein_id="BAA09593.1"
/db_xref="GI:1136633"
/translation="MTHDVLGAGLANGLIALLARAAPDLRVLLLDHAGSPDGH
WSCHDPLSDPDLARLKPLRRANWDOEVRFPFHARLRATGYSLDGAALADAVVRSG
AIRWDSDIALDAQATKLSCTRIEAGVLRGAQPSRLHVTGQFVGVIEIETR
PHGVPRMTMDATVQDQGYRITLIPSPTRILIEDTYSDDGDLDDALAAASHDY
ARQGGTGAERERERGIPIALAHDAAGFWADHAGPVPGLRAGFFHPTVGSYLYA
AQADVAGLSPGPDALRGATRDYAIRDRDRPRLNRLNRLFRGCPDRRYTILQ
RFRMPHGLIERFYAGRLSVADQLRVTKPPIPLGTALRCLPERPLLKENA"
2506. .4011
/gene="crtI"
2506. .4011
/gene="crtI"
/citation={2}
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="BAA09594.1"
/db_xref="GI:1136634"
/translation="MNAHSPAATAIVIGAGPGLALAILOSAGIATTLVBARDKPG
GRAYVNHQGLFDAGPTVITDDALKELWALTQDMARDVTLMVSPFRLMWPFGK
VFYVNEADQLEQIAQFNPDDEGFRFRDYABEYVQBYVKGITVPFKLQMLKA
APALMKLEAYKSHAKVATFKDYPYLRQAFVSHYTLVGNPNFSTSIYALNQLMKA
GVNFKGTQNLVAGWVALPERLQGMILNARVARIDTDPRTATVPLADGALATDM
VANSQDMHNYRDLIGHTARGQSRKSLNARWMSLFLVHLFGRLREAPKDVAAHTILF
GPYKELVNEIFKGLAEDFSLYLSHSPCTTDEMAPPGMSTHYVLAPVPHLGRADID
WAVEGRPHNDKTIIRNFYLVGAGTHFGAGIPGVVGSAAKATAQVMSDLASA"
4008. .4913
/gene="crtB"
4008. .4913
/gene="crtB"

/citation={2}
/codon_start=1
/transl_table=11
/product="Phytoene synthase"
/protein_id="BAA09595.1"
/db_xref="GI:1136635"
/translation="MSDLVLTSTEAITQGSQSFATAAKLMPGIRDDTVNLVANCRA
DDVIDGOALGSCPEAVNDPQARLDGLRADTLAALQGDGVPVTPFACARHDPFOAM
PMDLIRGFAMDEARVRLTDDVLEYSVHAGIVGVMMARVGMVRDVPDLDRACDLGL
AFQLTNIARDVIDDARVIRGSCYLPGLMDQAGARVDGVPSPELYTVILLLDAELVY
ASARVGLADLPPCRAWSIAAALRIYRAIGLRIRKGGPEAYRQRISTSKAAILGLIG
GWDVADHACRGRVATGLWTRPPHA"

ORIGIN

Query Match 14.8% Score 1037.2; DB 1; Length 5373;
Best Local Similarity 57.5% Pred. No. 1.6e-207;
Matches 2065; Conservative 0; Mismatches 1468; Indels 57; Gaps 9;
QY 2456 GCGTGCCGCCACATCGTCGAGCAGGCGCTGTGCGCAGCAGCAAGTCGTGCTGGCGGAGGC 2515
DB 1277 GCGCCCGTCGACAAGCTGAAGCAGGACCTGAAGATGTCGGGCGTGTGCGGCGCGAGGC 1336
QY 2516 GACCTGATGCGCAGCAATACGATGTGATTTGTGCGTGTCTGACCTGGCGAATGGCTTG 2575
DB 1337 GCAGAGCGCAGTGACCCCATGACGTGCTGTGCGAGGGGGCGCTTGGCCACGCGCTG 1396
QY 2576 ATTGCGCTGCGTCTGCGTCAATTGACGCCAACAATGCCCTGTGCTGTGCGAGAGCGAT 2635
DB 1397 ATCGCCCTGGCGTGGCGGGCGCGCCGACCTGCGCGTGTGCTGTGCTGGACCATGCC 1456
QY 2636 GCGCATCGGCGAGCAATCATCTCTGCTGCTTTTCATCAGAGCGATCTCAGCGCGCAACA 2695
DB 1457 GCAGGACCGTCAGACGCGCCACACCTGTGCTGCCACGACCCCGACCTGTGCCGCGACTGG 1516
QY 2696 CTTCGCTGGCTGCAACCGCTGATTACCGTGGTGTGCTCAGGTATTCAGGTTCGTTTCTT 2755
DB 1517 GTGCGCGGCTGAAGCCCTGCGCGCGCCCACTGGCCCGACGAGGAGTGGCGTTTCCC 1576
QY 2756 GCGTGGCGCCCAATCTGAGACGGGGATATTGTTCCATCGCATCAGGCGATTTGCCCGC 2815
DB 1577 CGCCATGCCCGCGCTGGCCACCGGTTACGGGTGCTGGACGGGGCGGCGCTGGC---G 1633
QY 2816 CATCTTTACGGCGGATGGGTGACGATCTGTGGCAACAACACAGCGCTACAACAGGTAAAA 2875
DB 1634 GATCGGTGGTCCCGTCCGGCCCGCAGATCCGCTGGGACAGCAGCATCGCCCTGCTGGAT 1693
QY 2876 CCCACGCAAGTGCAGCTGGCGGATGGCGGTGAACTTGTGTCGCAAGTGGTGAATGATGGT 2935
DB 1694 GCGCAGGGGGCGACGCTGCTCTGCGGCACCCGATCGAGGGCGGCGCGCTCTCGACGGG 1753
QY 2936 GCGGCGCTGCAGCCGACGCGCATCTGACGTGGGTATTCAGGTGTTTCTTTGGACAAGAG 2995
DB 1754 CGGGCGCGCAGCGCGCTGACCGGCGATCTGACCGGCGTTCAGAAATTCGTGGGTGTCAG 1813
QY 2996 TGGCAGCTGGCGCAGCGCGCTGCGAGCAGCGCATCTCTGATGGATGCCACCGTCCGAT 3055
DB 1814 ATCGAGACGACCGCGCCCGCAGCGCGTGCCTGCGCGCGATGATCATGACGCGACCGTACC 1873
QY 3056 CAGCAAGCGGGTATCGTTTGTCTACAGCTGCGCTCAGGCGCGATCGGCTATTGATT 3115
DB 1874 CAGCAGGACGGGTACCGCTTCTATCTGCTGCGCTTCTCTCGACGCGCATCTCTGATC 1933
QY 3116 GAAGATACCCATTACGTTAAACAGCCCGCTGGCGGAGAACACCGCTCGTCAGCACATC 3175
DB 1934 GAGGACACCGGCTATTCGATGGCGCGCATCTGGACGACGAGCGCTGGCGCGCGCTCC 1993
QY 3176 GCGGCACTATGCCAATCAGCAAGGCTGGACGCTGAGTACGCTGCTCGGTGAGAGACGCGC 3235
DB 1994 CAGCACTATGCCCGCAGCAGGCGCTGGACCGGGCGCGAGGTCCGCGCGAA---CGCGGC 2050
QY 3236 ATATTACCGATTACCTGAGCGGCAACATCGATCCATTCTGG---CAACAGCAGCGCGGC 3292
DB 2051 ATCTCTCCCATCGCGCTGGCCCATGATGCGGCGGGCTTCTGGGCGGATCACGCGCGGGG 2110

QY 3293 CAAGCGTGCAGCGGCTCGCGCGCGGCTGTTTCATGCGCACACCGGTTACTCTTGGCG 3352
DB 2111 CCTGTTCCCGTGGGACTGCGCGCGGGTCTTTTCATCCCGGTCAACCGGCTATTCTGCTGCC 2170
QY 3353 TCGCGCGTGGCGTACGAGTGTGTAGCAGCGCTGT---TGCCACCGATGCCCTCAGC 3409
DB 2171 TATGCGGCACAGGTGGCGGACGTTGTGGCGGGTCTGTGCGGGCGCGCGGACCGAGCGG 2230
QY 3410 CTCAGCAACATATCGAACGCTTTGCGCGTCAAGCATGTGCGGAAACAGCGATTTTTCGT 3469
DB 2231 CTGCGCGGCCATCGCGATTACGCGATCGACCGGGCGCGCGGACCGCTTCTGCGC 2290
QY 3470 CTGCTAAACCGCATGCTGTTTGGCCGTTAAGCCGACAGCGGCTGGCGCGTGTATGCAA 3529
DB 2291 CTTTGAACCGGATGCTGTTCCGCGGCTGCGCGCCGACCGCGGCTATACCTCTGCTGCAG 2350
QY 3530 CGTTTTTACCGGCTCGATCCCGGTTAAATTAGCCGCTTTTACGCGGGCAACTGCGCCTG 3589
DB 2351 CGGTTCTACCGATGCCGATGACTGATCGAACGGTTCATGCGCGCGGCTGAGCGTG 2410
QY 3590 CCGGATAAACCGGATTCTGTGCGCAAGCCCGGTCGCCATCGGTGAAGCGCTGCGC 3649
DB 2411 CGGGATCAGTGGCATCGTACCGCGCAAGCCTCCCATTTCCCTTGGCACGCGCATCCGC 2470
QY 3650 GCGCTGTGAATTCTGTGCAACGAGGGAAGAAATAAGAACGCACTT----- 3697
DB 2471 TGCCTGCCGAAACGTCCTGCTGAAGGAAACGCAATGAACGCCCATTCGCCCGCGGCCA 2530
QY 3698 -----ATGTGATTGGCGAGGCTTTGGGCGCTGGCGTGGGATTCGCTGCAAG 3748
DB 2531 AGACCGCATCGTGTACGGCGAGGCTTTGGCGGGCTGGCCCTTGGCCATCCGCTGCACT 2590
QY 3749 CGGCGGGCATACCAACCACTTACTCGAGCAGCGGACAAACCGGGCGGACGCGCTATG 3808
DB 2591 CGGCGGGCATCGCCACCACTGCTGAGCGCGCGGACAAAGCCCGGGCGCGGCTATG 2650
QY 3809 TGTTTGAAGACAGTGGCTTTACTTTCGATGCGCGGACCAACGGTGATCACCGATCCAGCG 3868
DB 2651 TCTGGCACGATCAGGCGCATCTCTTCGACGCGGGCGGACCGCTCATACCGACCCCGATG 2710
QY 3869 CCATCGAAGAGTGTTCACGCTGGCAGGAAATTCGCTACGAGCATTCGTCGAGCTGATGC 3928
DB 2711 CGCTGAAGAGTGTGGGCGCTTGAACCGGCGAGCATGCGGCGGACGTGACGCTGATGC 2770
QY 3929 CGGTAAACGCTTCTATCGCTGTGTGGGAAAGTGGCAACAGCTTGTATTACGACAATA 3988
DB 2771 CGGTCTCGCCCTTCTATCGGCTGATGTGCGCGGGGGAAGGTCTTCGATTAGGTGAACG 2830
QY 3989 ATCAGCGCGTGTGAGGACAGATGCGCACGTTTCAATCCGCAAGATGTAGAAGGCTATC 4048
DB 2831 AGGCGGATCAGCTGGAAACGCCAGATCGCCCAAGTTCAACCCGACGACCTTGGAGGATACC 2890
QY 4049 GTCAATTTCTGCTATTTCACGTGAAGTATTTAGAGAGGTTTATCTGAACTCGCACGG 4108
DB 2891 GCGGCTTCCGTATACGCTGAGAGGTGTACAGAGGGGCTACGTTCAAGCTGGGACCG 2950
QY 4109 TGCGGTTTCTGAGGTGCGTACATGCTCGCGGTGCGCGCGGAGTTGGGACGCTCTGCAAG 4168
DB 2951 TGCGCTTCTCAAGCTGGGCGAGATGCTCAAGCGCGCGCGCTGATGAAGTTGGAG 3010
QY 4169 CATGGCGAGCGTCTACAGATGGTGGCGAAATTTATTCAGGACGATCATCTGCGTCAAG 4228
DB 3011 CCTCAAGTGGTCCATGTCGAAGGTGCGCACCTTTCATCAAGGACCCCTTATCTGCGGCGAG 3070
QY 4229 CGTTTTTCTTCCACTTATCTGTTGGGCGGTAATCTTTTTCGAAGTCAATGATCTATA 4288
DB 3071 CGTTTTTCTGATCACAGCTGCTGTTGGGCGGGAATCCCTTCTCGACCACTGATCTATG 3130
QY 4289 CTTTAATTCATGCGCTGAGCGTGAATGGGGGCTGTGTGTTTCCGCGCGCGGACCGCGG 4348
DB 3131 CGCTGAACCAACGCGCTGGAGCGGCGCGGGGTCTGGTTTCGCAAGGGCGGACCAACCC 3190

QY 4349 CGCTGGTGCAGGGCATGGCGCGACTGTTTCAGGACTTGGGCGCGAGCTGTTACTGAATG 4408
DB 3191 AGCTGGTTGCGGGCATGTGTCGCGCTGTTTCGAAACGGCTTGGCGGGCAGATGCTGCTAAACG 3250
QY 4409 CCGAAGTGAAGCAGCTGGAAACAGCGGCAATTCGATTAAGCGCGTTCAGTTAGAGGGCG 4468
DB 3251 CCAAGGTTCGCGGGATCGACACGGAGGGCGCGCGGACCGCGGTGACCTTGCGCGCAGC 3310
QY 4469 GACGACGCTTCGATCCGCGCTGTGGGCTTCAATTCGCGACGTGTGTCATACCTACGACA 4528
DB 3311 GCGCGCCTTGACCGCGGACATGGTCGCGCAGCAAACCGCGACGTGATGACAACTATCGCG 3370
QY 4529 AACTGCTTTCGCCACCATCCGCTGGCAATGAAACGTCGACATCGCTGAAGCGTGAAGCGCA 4588
DB 3371 ACCTCTGGGCGCATACCGCGCGCGGACAGCGCGGCGGAAATCGCTGAAACGCGAAGCGT 3430
QY 4589 TGAGCAACTCGCTGTTGTTGACTCTATTTTGGGCTCAATTCAGCGCGCATGAACAGCTCGCG 4648
DB 3431 GGTCCATGTGCTCTTTCGTGCTGCAATTTTCGGCTTGGCGAGGGGCCCAAGGACGTGGCG 3490
QY 4649 ACCACACCGTCTGTTTTCGCGCGGTTATCTGTAGTTGATCGATGAGATTTTTCACAGCA 4708
DB 3491 ATCACACCATCTGTTTCGCGCGCGCTACAAGGAGCTGGTCAACGAGATCTTCAAGGGCG 3550
QY 4709 GCCAGCTGGCAGACGATTTTTCACCTTACCTGCAACCGCGCTGCGACGACGATCCGTCGC 4768
DB 3551 CGAAGCTGGCGAGGATTTCTCGCTCTATCTTCATTCGCGCTGCAACGACCGCGGAGA 3610
QY 4769 TGGCACCGCGCGCTCGGCGAGCTTTTATGTGTAGCGCGGCTGCGGACCTCTCGGCACCG 4828
DB 3611 TGGCGCTTCCGGGATGTCACGCAATYACGCTCTGCGCGCGGTGCGGATCTTGGGCGCG 3670
QY 4829 CTGACATCGACTGGCAACAGGAAGACCGCGCTTGGCGATCGAATTTTTCGTTTACTG 4888
DB 3671 CGGACATTTGTTGGCGGTTCGAGGGCGCGGCTATGCGGACCGCATCTCTGGCGAGCTG 3730
QY 4889 AGCAGCACTACATGCGGGATTCAGTCAGCAATTAGTGACACAGAAATGTTTACGCCCT 4948
DB 3731 AGGAGCGGCTGATCCCGAACCTTGCAGCGCAACTGACCAACGACGCGCATCTTACCGCCG 3790
QY 4949 TTGATTTTCGCGACACGCTGATGCCATCACGGCTCGGCGTTCGCTGGAGCGCGATTT 5008
DB 3791 CCGATTTTCGCGACGAACTGAACGCCCATTCATGCGAGCGCTTCTCGTTCGAGCGCATC 3850
QY 5009 TGACCAAGCGCTGTTTCGCGCGCATAAACCGGATGCGGATATCAGCAATCTCTATC 5068
DB 3851 TGACGCAATTCGCGTGGTTTCGCGCGCAACACCGGCAAGACGATCCGCAACTCTATC 3910
QY 5069 TGGTGGGTGCGGTACGATCCAGGGCGGGCGTGGCGCGGCTGATCGGTTGCGGCAAG 5128
DB 3911 TGGTGGGCGCGGCAACCCATCCGGCGGGGGAATTCGGGGCGTGTGGGCTCGGCGAAG 3970
QY 5129 CCACCGCGAGGCTGATGCTGGAGGATCGCGCGCAATGAATCGACAGCTTTTACTTGAGCA 5188
DB 3971 CCACCGCGCAGTGTATGCTGCTCGACCTTGGCGAGCGCATGAGCGATCTGGTCTT---GAC 4027
QY 5189 AGTAACGCAAAACCATGGCGGTGGCTCGAAGAGTTTCGCCACCGCGCGCAAGCTGTTGA 5248
DB 4028 CTCGACCGAGGCGATCACCAAGGTCGCAAGCTTTCGCCACCGGCGCCAAAGCTGATGCC 4087
QY 5249 TGCAACGACCGCGCGGACGCTGATGCTGTATGGTGGTGTGCTCACTGCGATGATGT 5308
DB 4088 GCGGGGCACTCCGCGACGACACGCTGATGCTCTATGCTGTTGCTGCGGCGGCGGATGACGT 4147
QY 5309 GATTGATGGGCAAAACGCTGGGCGAAGGCGGACGAGCATGCGCTCGAAGAGACGCGCAGGC 5368
DB 4148 GATCGACGCTAGGCGCTTG-----GGCAGCTGCCCGAGGGCGGTGAACGACCGGAGGC 4201
QY 5369 ACGTATGACGATCTGCAAAATGAAACCGCGCGCTTACAGCGGCGGCGCAATGATGA 5428
DB 4202 GCGGCTGAGCGGCTGCGCGCGACACGCTGCGCGGTTGACAGGGCGACGGTCCGCTGAC 4261
QY 5429 ACCGCGTTTAGGGCGTTTCAGGAAGTGGCGATCATTCACAGCTGCGCGCAACACTGGC 5488

```
Db 4262 CCCGCCCTTTGC-----CTGCGCGGTGCGCGCGCATGATTTTCGCGAGGCGCTGGCC 4315
QY 5489 GTTTCATCATCTGGAAGGCTTCGCTATGATGACGCAACGAACATTACGCGAGCTTCGA 5548
Db 4316 CATGACCTGATCGAAGGCTTTGCGATGATGTCGAGGCGCGGACTATCGACGCTGGA 4375
QY 5549 TGACACGCTGCGTTACTGCTATCAGTTCGCGGGCGTGTGCTGTTGATGATGCGCGCGGT 5608
Db 4376 CGACGTGCTGGAATATTCCTATCAGCTCGCGGCGATCGTGGCGGTGATGATGGCCCGGT 4435
QY 5609 AATGCGCTGCGCGACGAAGCGTCTCGATCACCGCTCGCATTTAGGACTGCGGCTTCCA 5668
Db 4436 GATGGCGGTGCGCGACGATCTCTGCTGGAACGCTCGGACCTGGGCGTGGCGTTCGA 4495
QY 5669 GCTCACTAATATGCGCGCGACATTTGTAAGATGCGGAATGCTGCTGCTATCTGCC 5728
Db 4496 GCTGACCAACATCGCGCGCGATGATCGACGCGCGCATCGCGCGGTGCTATCTGCC 4555
QY 5729 GAATCTGCTGCTGATCAGCGCGGATTAGCGCGCATAGCTGACGCGCACCGACATCG 5788
Db 4556 AGCGCACTGGCTGATCAGCGCGCGCGCGGTGCGCGCGCGTTCGCGCGTTCG 4608
QY 5789 TGCAGCGCTCGCTCACTCGCAGCGCTTTAGTGGCGGAGCGCGAACCTATTATCACTC 5848
Db 4609 --CGGAGCTGTACACCGTATCTCGCGCTGCTGGATCGCGTGAATCTATTACGCGTC 4666
QY 5849 GCGCGCATCGGTTTACCGGTTTACCGTGGCTGCGGCGGCGGCGGCGGCGGCGGCTGCG 5908
Db 4667 GCGCGCGGTGCGTCTGCGCGGATCTGCGCGCGGCTGCGCGTGGTCCATCGCGCGCGGT 4726
QY 5909 CGCGGTTTATCGGAATTCGGGTCAAGCTTCAGCAGCGCGGTGTCAGCGCTGGGATTC 5968
Db 4727 CGGATCTATCGCGCATCGGCTGCGCATCGCAAGCGCGCGCGGCGGCGGCTATCGCCA 4786
QY 5969 ACAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 6018
Db 4787 GCGGATCAGCAGCGCTCAAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4836

RESULT 14
FAU62808/c 5188 bp DNA linear BCT 18-FEB-1997
LOCUS Flavobacterium ATCC21588 geranylgeranyl synthase (crtE), phytoene
DEFINITION synthase (crtB), phytoene desaturase (crtI), lycopene cyclase
(crtY), b-carotene hydroxylase (crtZ) genes, complete cds.
ACCESSION U62808
VERSION U62808.1 GI:1842241
KEYWORDS
SOURCE Flavobacterium sp. ATCC 21588
ORGANISM Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
Flavobacteriaceae; Flavobacterium.
REFERENCE 1 (bases 1 to 5188)
AUTHORS Pasamontes,L., Hug,D., Tessier,M., Hohmann,H.P., Schierle,J. and
van Loon,A.P.
TITLE Isolation and characterization of the carotenoid biosynthesis genes
of Flavobacterium sp. strain RL534
JOURNAL Gene 185 (1), 35-41 (1997)
MEDLINE 91786694
PUBMED 9034310
REFERENCE 2 (bases 1 to 5188)
AUTHORS Pasamontes,L.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1996) VFCB, F. Hoffmann-La Roche Ltd., Basel,
4070, Switzerland
FEATURES
source location/Qualifiers
1..5188
/organism="Flavobacterium sp. ATCC 21588"
/mol_type="genomic DNA"
/strain="RL534"
/db_xref="taxon:50286"
95..982
gene
```

```

/gene="crtE"
95..982
/gene="crtE"
/codon_start=1
/transl_table=11
/product="geranylgeranyl synthase"
/protein_id="AAC44848.1"
/db_xref="GI:1842242"
/translation="MTPKQFFPLRLDLVEIRLAQISQQFGVVSAPLGAAMSDAALSPGK
RFAVLMLVAESSGGVCDAMDVAACAVEMVHAASLI FDDMPCDDMTTRRGQPATVH
AHGEGRAVLGALITTEMRILLEGARGATDQARLIVASMSRAMGPVGLGACAGDLH
APKDAKIERQLKTVL FVAGLEMLSIIKGLDKAETQLMAFGRLGCVFOSYDDL
LDVIGKASTGKDTARDTAPGPKGLMAVGMDVAQHYRASRAQLDELDELMTFLFRG
GQIADLLARYLPHDIRSA"
complement(979..1890)
/gene="crtB"
complement(979..1890)
/gene="crtB"
/codon_start=1
/transl_table=11
/product="phytoene synthase"
/protein_id="AAC44849.1"
/db_xref="GI:1842243"
/translation="MTDLTATSEAAIAQGSQSFQAQAKLMPGIRREDTVMVLYAWCRHA
DDVIDGVMGSAFEGAGDPOARLICALRADTLAALHEDGPMSPFALROVAREHDPD
LWPMDLTEGFAMDVAADREYRSKSLDDVLEIYTHVAGVGVMMARVWGVDVADLRACDL
GLAFQLTNIARDVIDDAIRGCYLPADLAEAGATVEGVPVSADLYSLIRLLDAAEP
YASARQGLPHLPKPCAWSIAAALRIYRAITGTRIQQGPEAYRQRISTSKAAKIGLLA
RGGIDAAASRLRGGEISRDGLWTRPA"
complement(1887..3371)
/gene="crtI"
complement(1887..3371)
/gene="crtI"
/codon_start=1
/transl_table=11
/product="phytoene desaturase"
/protein_id="AAC44850.1"
/db_xref="GI:1842244"
/translation="MSSAIVIGAGFGGLAIALRIQSAGIATTIVEARKDGGRAVYVN
DQGHVFDAGPTVTPDPSLRELWALSQPMERDVLLPVPFFRYLTWADGSRFFYVND
DDELIROVAGFNADVDGRRFHDYAEVREGLYKLTPTFLKQMLNAPALMRL
QAYRSVHMVARFIQDPLHQAFSLTLVGGNPFSTSIYALIHALLERGGVWFAGK
GTNOLVAGMALPERLGTLLINARVTRIDTEGDRATGCTVLLDGRALERTVNSGDV
MHSYRDLHGTRGRTKAAILNORWMSLFLVHFGLSKRPENLAHSVIFGPYKGL
MSEIFNGLRPDDFSNMLHSPCTVDPSLAPGEMSTHYVLAAPVPHLGRADVDEAEPG
YAEIRFLEERRAIPDLRKHILTVSRIDFSPADFSTELSAHGSASFVSPILTQSAWFRP
HNRDAIPNFYIVGAGTHPGGIFGVVGSAKATAQVWMLSLAVA"
complement(3368..4516)
/gene="crtY"
complement(3368..4516)
/gene="crtY"
/codon_start=1
/transl_table=11
/product="lycopene cyclase"
/protein_id="AAC44851.1"
/db_xref="GI:1842245"
/translation="MSHDLITAGLISGALIALAVRDRRPDARIWMLDARGSPSDQHT
WSCHTDIUSPEWLAKLSPIRGENTDQVAFDPHSRLTTCVGSIEAGALIGLQGVND
LRWNTHVATDDTGTATLDGSGRIEACACVIDARGAVETPHLTGFKQKFGVIEITDAPH
GVERPMINDATVPMQDGYRFTYLLPFSPTRILIEDTRYSDGDDDDGALQAASLDYAA
RRGTQGMERRERGIPLTALAHAIDAFWRDHAQGAVPVGLGAGLPHVPTGYSLSYAAQ
VADATAARDLTATASARRAVRWAIADRDRFLRLNMLFRGCPDPRRYLLQRFYR
LPQPLIERFVAGRLTLADRLRIVTGRPPILPSQAVRCLPERPLLQERA"
complement(4513..5022)
/gene="crtZ"
complement(4513..5022)
/gene="crtZ"
/codon_start=1
/transl_table=11
/product="b-carotene hydroxylase"
/protein_id="AAC44852.1"
/db_xref="GI:1842246"
/translation="MSTWAAAILTVILTVAAMELTAYSVHRWIMHGFLGWGWHKSHHDE
```

DHDALEKNDLYGVLPVAVISIVLPAIGAMGSDLAHWLAVGVTCYGLIYYFLHDLVHG
RMPFRVPRGVLRRVYQAHRMHVAHVHRENCVSFGFIWAPSVDSLKAEKRSGLLKK
DREGADRT"

ORIGIN

Query Match 14.8%; Score 1035.4; DB 1; Length 5188;
Best Local Similarity 56.5%; Pred. No. 3.8e-207; Mismatches 1536; Indels 30; Gaps 5;
Matches 2037; Conservative 0;

QY 2486 GTCCAGCAGCAAGTCTGCTGGCGAGCGGACCTGATCGCACGCAATACGATGTGATT 2545
DB GGGCGGCTGCTGAAGGACCGGAGGGCGGATCGCAATACATGAGCCATGATCTGCTG 4499
QY 2546 TTGTCGGTGTGGAAGTGGCTTGAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2605
DB ATCGCGGGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4439
QY 2606 CAACATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2665
DB GATGCGCGCATCGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4379
QY 2666 TTTCATCACAGCGATCTCAGCGCCGAACTTCTGCTGCTGCTGCTGCTGCTGCTG 2725
DB TGCCACGACACGATCTTTCGCGCCGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4319
QY 2726 CTTTGGTCAAGTATCAGTGGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2785
DB GAATGACAGCATCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4259
QY 2786 TGTTCATCGATCAGCGCATTTTTCGCGCCATCTTTACGCGCGCATGCTGCTGCTG 2845
DB GGCTCGATCAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4208
QY 2846 TGGCAAAACACAGCGCTCAACAGGTAAACCCAGCAGTGAAGTGGCGGATGGCGCT 2905
DB CGTGGATACGATGCTGCGAGCTTGGAGCATACCGCGCGAGCTGACGAGCGCTCG 4148
QY 2906 GAACTTGTGCGCAAGTGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2965
DB CGATGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4088
QY 2966 CTGGGTATCAGTGTCTTCTGCAAGAGTGGCAGCTGGCGCGCATGCTGCTGCTG 3025
DB GTGGGTTCAGAAATTCGTGGCGCTGAGATCGAGCATGAGCGCGCCCATCGCGCTG 4028
QY 3026 CAGCGCATCTGATGATGCCACCGCTGATCAGCAAGCGGGTTATCGTTTGTCTACAG 3085
DB CGCCCGATGATCATGACCGGACCGTTCCGAGATGAGCGGGTACCGCTTCTATCTG 3968
QY 3086 CTGCGCTCAGCGCGATCGGCTATTGATTGAAGTACCCATTAAGTTAACCGCCCGG 3145
DB CTGCGCTTCACTCCACCGCATCTGATCGAGGATACGCGCTACAGCGACGCGCGCAT 3908
QY 3146 TGGCGGAGAACACCGCTGTCAGCACATCGCGCATATCCCATATACCGTGAAGCG 3205
DB CTGAGATGCGGCGGAGAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3848
QY 3206 CTGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3265
DB GGGCAGGAGATGGGCGGAGAGGAGGATCTCTGCCCATCGCTGGCCCATGACCCATA 3788
QY 3266 GATCGATTCGGAACACAGCGCGGCGCAAGCGTGCAGCGGCTGCGCGCGCGGCTGTTT 3325
DB GGGCTTCTGCGGAGCACACGCGCAGGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3728
QY 3326 CATGCCACCGGTTTACTCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3385
DB CACCCGTCACCGGATATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3668
QY 3386 CTGTTGCCACCGGATGCCCTCAGCTCAGCAACATATCGAAGCTTTTCCCGCTCAGCAG 3445
DB CGCGACCTGACGACCGCTCGCGCGCTGCGCGGCTGCGCGGCTGCGCGCATCGATCGCG 3608

QY 3446 TGGCGGCAACAGCGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3505
DB GATCGCG---ACCGCTTCTGCGGCTGCTGAAACCGGATGCTGTTTCGCGGCTGCGCGCC 3551
QY 3506 CAGCAGCGCTGGCGGTGATGCAACGTTTATACCGGCTGATGCTGCGGGTAAATATGACCGC 3565
DB GACCGTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3491
QY 3566 TTTTACGCGGGCAACTGCGCTGCGCGATAAAAACGCGGATTTCTGTCGCGCAAGCGCGC 3625
DB TTTATGCGGGCGCTGACATGTCGCGACCGGCTTGCATGCTCAGCGACCGCGCGCC 3431
QY 3626 GTGCCCATCGGTGAAGCGCTGCGCGCTGTTGAATCTGTCGAAACCGAGGAGAGAAAA 3685
DB ATTCGCTGTCGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3371
QY 3686 TGAACACGACATTTATGATTTGGCGCAGCGCTTTGGCGGCTGCGGCTGCGGATTCGCTG 3745
DB TGAGTTCCGCCATCGTCACTGCGCGCAGGTTTCGCGGGGCTTGGCTTGCATCCGCTG 3311
QY 3746 AAGCGCGGGCATACCAACCTTACTGAGCAGCGGCAAAAACGCGGCGGACGCGCT 3805
DB AATCGCGGCACTCGGACCACTGCTGAGGCGCGGACAAAGCGCGCGCGCGCT 3251
QY 3806 ATGCTTTGAGCAGAGTGGCTTTACTTTCGATGCGCGGACCGGCTGATCACCGATCCCA 3865
DB ATGCTGGAACGATCAGGGCCACGCTTTTCGATGCGAGGCGCGGCTGCTGCTGCTGCTG 3191
QY 3866 GGGCATCGAAGAGTTTTCAGCTGGCAGGAAATTCGCTCAGCGATTAAGTCTGAGCTG 3925
DB ACAGCTCGGAGAGCTTGGGCCCTCAGCGGCAACCGATGAGCGCTGCTGCTGCTGCTG 3131
QY 3926 TGCGGTAAACGCGCTTCTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3985
DB TGCGGTCTCGCGCTTCTACCGCTGACATGCGCGGACGCGCGCGAGCTTCGAATACG 3071
QY 3986 ATAATCAGCGCTGCTGAGCAGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4045
DB ACGACGACGAGCTGATTCGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3011
QY 4046 ATCTGCAATTTCTGCTTATTCAGCTGAAGTATTTAGAGAGGTTTCTGAAACTCGCA 4105
DB ATCGCGGCTTCCAGATTAACGCGAGGAGTCTATCGCGAGGCTGCTGCTGCTGCTGCTG 2951
QY 4106 CGGTGCGCTTCTGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4165
DB CCACGCGCTTCTGAAAGTGGCGCAGTCTGAAACGCGCGCGCGCTGCTGCTGCTGCTG 2891
QY 4166 AAGCATGGGCGAGCTTACAGCATGCTGCGGCAAAATTTATTCAGGACGATCATCTGCTG 4225
DB AGGCATACCGCTCGGTTCACAGCATGCTGCGCGCTTCTATCCAGGACCGCATCTGCGCG 2831
QY 4226 AGGCGTTTCTTCCACTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4285
DB AGGCTTCTGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2771
QY 4286 ATACCTTAATCATCGCTGAGCGTGAATGCGGCGTGTGCTGCTGCTGCTGCTGCTGCTG 4345
DB ATGCTGCTGATCATCGCTGGAACGCGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2711
QY 4346 GGGCGCTGTCAGGCGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4405
DB ACCAGCTGCTGCGGGCATGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2651
QY 4406 ATGCGCAAGTGAAGCAGCTGGAACACGCGCAATCGCATTAAGCGCGTTCAGTTAGAGG 4465
DB ATGCGCGCTGTCAGCGGATCGACACGAGGCGGATCGCGGCGGCTGCTGCTGCTGCTG 2591
QY 4466 GCGGACGAGCTTTCGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4525
DB ACGGGCGGAGTTGCGCGCGGATACGCTGCGGCAACGCGCGAGCTGATGCAAGCTATC 2531
QY 4526 ACAAACTGCTTCGCCACCATCCGCTGGGCAATGAAAGCTGCGCATCGCTGAAGCGTAAGC 4585

Db	2530	GCACCTGCTGGGCCATACCCGCGCGGCGCACCAAGCGCGGATCCTGAAACCGGACG	2471
Qy	4586	GCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCTCTGATCAGCCGCGATGACACGCTCG	4645
Db	2470	GCTGGTCGATGTCGCTGTTCTGTCGATTTGGCCTGTCCAAAGCGCCCCGAGAACTCG	2411
Qy	4646	CGCACACACCGTCTGTTTGGCCCGCGTTATCGTGAGTTGATCGATGAGATTTTCAACA	4705
Db	2410	CCACACACGCTATCTTCGGCCCGGCTACAAGGGGTGGTGAACGAGATCTTCAACG	2351
Qy	4706	GCAGCCAGCTGGCAGACGATTTTTCATCTTACCTGCAACGCGCCCTGAGCAGGATCCGT	4765
Db	2350	GGCCAGCCTGCGGACGATTTCTCGATGTATCTGCAATTCGCCCTCGTGACGATCCCA	2291
Qy	4766	CGCTGGCACCGCCGCTGCGGAGCTTTTATCTGTTAGCGCGGTGCGGCATCTCGGCA	4825
Db	2290	GCCTGGCCCCCGAGGGGATGTCACGANTACGTCCTTGGCCCGTTTCGCATCTGGGCC	2231
Qy	4826	CCGCTGACATCGACTGCGAACAGAGAGAACCGGCTTCGCGCATCGAATTTTGTCTTATC	4885
Db	2230	GGCCGATGTGATTGGAAAGCGAGSCCGGGCTATGCCGAGCGCATCTTCGAGGAAC	2171
Qy	4886	TGAGCAGCACTATACGTCGCGGATTCGTCAGCAATTAGTGTACACACAGAAATGTTACGC	4945
Db	2170	TGGAGCCCGCGCCATCCCGACCTGCGCAAGCACCTGACCGTCAGCCGATCTTCAGCC	2111
Qy	4946	CGTTTGATTTTTCGCGACACGCTGCATGCCCATCACGGCTCGGGCTTTTCCTCGAGCCGA	5005
Db	2110	CCGCCGATTTACGACACCGAACTGCGGCCCATCAGCGCAGGCTTCTCGTCGAGCCGA	2051
Qy	5006	TTTTGACGCAAGCGCTGCTTCGCGCCGCAATACCGCGATGCCGATCAGCAATCTCT	5065
Db	2050	TCCTGACGCAATCCGCTGTTTCGCGCGCATAACCGCGACCGCGCATCCCGAACTTCT	1991
Qy	5066	ATCTGTTGGGTGCGGTTACGCATCCAGCCGCGGCGTCCCGCGTGATCGGTTTCGCGCA	5125
Db	1990	ACATCTGGGGGCGGGCAGCATCCGGTTCGGGCAATCCCGGTGTGTTGGCAGCGCCA	1931
Qy	5126	AGGCCACCGCAGGCTGATGCTGAGAGATCGCGCGGAATGAATCGACAGGCTTTACTTGA	5185
Db	1930	AGGCCACCGCGCAGGTCATGCTGCGACCTTGGCCGTCGATGACCGATCTGACGGCGAC	1871
Qy	5186	GCAAGTACGCAAAACATGCGGCTGGCTCGAAGAGTTTCCGACACCGCGCCAGCTGTT	5245
Db	1870	TTCCGAAGCG---GCCATCGCGCAGGGTTGCAAAAGCTTCGCGAGCGCGCAAGCTGAT	1814
Qy	5246	TGATGACCGCAGCGCGCAGCAGCTGATGCTGTGCTGGTGTGTCCTCACTGCGATGA	5305
Db	1813	GCCGCCGGCATTCGCGAGGATACGGTCATGCTCTATGCTGTTGCGAGCATGCGGATGA	1754
Qy	5306	TGTGATTGATGGCAAAACGCTGGGCGAAGGCGGACGCGAGCATCCGCTCGAAGACGCGCA	5365
Db	1753	CGTGATCGACGGCAGGTGATGGGTCTTCGCCCGCGAG-----CGGGCGCGCGACCCACA	1700
Qy	5366	GGCAGTATGACGATCTGCAAAATTGAAACCCGCGCGCTACAGCGCGCGCATGGA	5425
Db	1699	GGCGCGCTGGGGCGCTTCGCGCGCGACACGCTGGCGCGTGCACAGAGAGGCCCGAT	1640
Qy	5426	TGAACCGCGGTTTAGGGCGTTTCAGGAAGTGGCGATCATTCACAGCTGCCGCAACACT	5485
Db	1639	GTCCGCGCCTTCGCGCGCTGCGCAGTCCGCGGGCATGATTTCCCGGACCTTTG	1580
Qy	5486	GGCGTTTGATCATCTGGAAGGCTTCGCTATGTGATGCACGCAACGAACTATACGCGACTT	5545
Db	1579	GCCGATGGAACCTGATTCGAGGGTTTCGCGATGATGATGTCGCGATCCGCAATACCGAGCCT	1520
Qy	5546	CGATCACACGCTGCTTACTGCTATCACGTCGCGGGCGTGTTCGCTTTGATGATGGCGCG	5605
Db	1519	GGATGACGTGCTGGAAATATCTTACCACTCGCGGGGTCTGGGGCGTGATGATGGCGCG	1460
Qy	5606	CGTAATGGGCGTGCAGCAAGACGGGTGCTGCATCACGCTTCGATTTAGACTTGGGCTT	5655

Db	1459	GGTGTATGGGCGTGCAGGACGATCGGCTGTGGATCGGCCCTGCGATCTGGGCTTGGGCTT	1400
Qy	5666	CCAGCTCACTAAACATTGCGCGCAGACATTTGAGAAGATGCCGAAAAATGGTCGCTGCTATCT	5725
Db	1399	CCAGCTGACGAAACATCGCTCGCGACGCTGATCGACGATGCCCATCGGGCGCTGCTATCT	1340
Qy	5726	GCCECAATCCTGGCTCGATCAGCGGGGATTTACGGCGCGGATACGCTGACTGCACCGCAACA	5785
Db	1339	GCCTGCCGACTGGCTGGCCGAGCGGG-----GGGACCGGTTGAGGGTCCGGTGCC	1289
Qy	5786	TCGTGCAGCGCTCGCCCTCACTGGCAGCGCGTTTAGTGGCGGAGCGGAAACCTTATTATCA	5845
Db	1288	TTGCGAGCGCGCTTATTTCGCTCATCCGCTGCTTGACGCGCGCAGCCCTATTATGC	1229
Qy	5846	CTCGCGCGCGATCCGGTTTACCGGGTTTACCGCTGCGCTCGCGTGGGCCATTCGCTACGGC	5905
Db	1228	CTCGCGCGCGCAGGGGCTTCGCACTTCGCGCGCGCGTGCGCCTGCTGATCGCGCGCGC	1169
Qy	5906	TCGGGCGGTTTATCGCGAAATTTGGCGTCAAAAGTTCAGCACGCCGCTGTGCACGCTCGGA	5965
Db	1168	GCTGCGTATCTATCGCGCAATTCGGGACGCGCATCCGCGAGGTGGCCCCGAGGCTATCG	1109
Qy	5966	TTACCGCGCAGCGCACCAAGTAGTGAAGGTGAAAACTGGCGCTCTGTTGTAAGGGCGAGTTT	6025
Db	1108	CCAGCGGATCAGCACGTCGAAGCTGCCAAGATCGGCTTCTGCGCGCGGAGGCTTGGGA	1049
Qy	6026	GGCGATCATCTTCGGCTGTGTCTGCTCTGAACCGCGTCCGCTGGTCTGTGGCAGCGTCC	6085
Db	1048	CGCGCGCGCATCGCGCTCGCGCGCGCGGCAAAATCAGCGCGACGGCCTGTGGACCCGACC	989
Qy	6086	TCG 6088	
Db	988	GCG 986	
RESULT 15			
AR169831/c			
LOCUS	AR169831	8625 bp	DNA linear PAT 17-DEC-2001
DEFINITION	Sequence 1 from patent US 6291204.		
ACCESSION	AR169831		
VERSION	AR169831.1	GI:17907751	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 8625)		
AUTHORS	Pasamontes, L. and Ieygankov, Y.		
TITLE	Fermentative carotenoid production		
JOURNAL	Patent: US 6291204-A 1 18-SEP-2001;		
FEATURES	Location/Qualifiers		
source	1..8625		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match	14.8%;	Score 1035.4;	DB 6; Length 8625;
Best Local Similarity	56.5%;	Pred. No. 3.8e-207;	
Matches 2037;	Conservative	0; Mismatches 1536;	Indels 30; Gaps 5
Qy	2486	GTGCCAGACGCAAGTCGTGTGTCGGCGGAGCGGACCTGATGCCGACCAATACGATCGTATT	254
Db	6984	GGCGCGCTGCTGAAGGACCCGCAAGGGCGGATCGCAATATATGAGCCATGATCTGCTG	6922
Qy	2546	TTGTCGTGCTGGACTGGCGAAATGGCTTGATTTGGCTGCGCTCTCGCTCAATTGACGCCA	2601
Db	6924	ATCGGGGCGGGGCTGTCCGGTGCCTGATTCGCGCTTGCCGTTGCGGACCGCAGACCG	6866
Qy	2606	CAACTGAAATGCTGTTGCTGGAGACGATGGCGATCCGGCAGGCAATCATACCTGCTGCG	2666
Db	6864	GATGCGCGCATCGTGTATGCTCGACGCGCGTCCGGCCCTCGGAACGACACCTGGTGCC	6801
Qy	2666	TTTTCATCAGCGATCTCAGCGCCGAAACAACTTCGTGCTGCTGCAACCGCTGATTACCGTG	2721

Db	5804	TGCCACGACACGGATCTTTTCGCCCGAATGGCTGGCGGCGCTGTGCGCCATTCTGTCGGCG	6745
Qy	2726	CGTTGGTCAAGTTATCAAGTGGCGTTTCTGCGCTGCGCGCCGAACTCTGACCGGGATTAT	2785
Db	6744	GAATGGACGGATCAGAGGTCGGTTTCCGACCAATTCGCGCGCTGACGACAGGCTAT	6685
Qy	2786	TGTTCCATCGCATCAGGGAATTTTCCCGCCATCTTTACGCGGCGATGGTGAAGATCTG	2845
Db	6684	GGCTCGATCGAGCGGCGCGCTGATCGGGCTGCTGCA-----GGGTTCGATCTG	6634
Qy	2846	TGGCAAAACACAGCCGTAACACAGGTAAACCCACGACAGTGCAGCTGGCGGATGGCCGT	2905
Db	6633	CGGTGGATACGATGTCGAGCGTGGACGATACCGGCGGACGCTGACGAGCGGCTCG	6574
Qy	2906	GAATTCGTGCGCAAGTGGTGAATGATGTCGCGGCTGCAAGCCGACGACATCTGCAG	2965
Db	6573	CGGATCGAGGCTGCCTGCGTATCGACGCCGCTGGTGGCGTGCAGACCCCGCACCTGACC	6514
Qy	2966	CTGGGTTATCAGGTGTTCTTGGACNAGATGGCAGCTGGCGGACGCGCCGACCGCTGCAG	3025
Db	6513	GTGGGTTTCCAGAAATTCGTGGCGCTCGAGATCGAGACCGACGCGCCCGCATGGCGTCGAG	6454
Qy	3026	CAGCCGATCCTGATGATGCCACCGTCGATCAGCAAGCGGTTATCGTTTGTCTACACG	3085
Db	6453	CGCCCGATGATCATGGACGCGACCGTTCCGAGATGAGCGGTAACGCTTCACTATCTG	6394
Qy	3086	CTGCCGCTCAGCGCCGATCGGCTATTGATTGAAGATACCCATTACGTTAACAGCCCGCG	3145
Db	6393	CTGCCCTTCAGTCCCAACCGCATCTGATCGAGGATAGCGCTACAGCGACGCGCGCAT	6334
Qy	3146	CTGGCGGAGAACACCGCTCGTCAGCNACATCGCGGACTATGCCAATCAGCAAGCGCTGAGC	3205
Db	6333	CTGAGCATGGGCGCTGGCGCAGGCGTGCCTGGGATATGCGCCGACGCGGGGCTGGACC	6274
Qy	3206	CTGAGTACGCTGCTGCTCAAGACGACGCGCATATTACCGATTACCTGAGCGGCAACATC	3265
Db	6273	GGGCGAGGATGCGGCGGAAAGGGGCATCTGCCCATCGCGCTGGCCCATGACGCCATA	6214
Qy	3266	GATCGATTCTGCAACAGACGCGCGCCGCAAGCTGTCAGCGGCTGCGCGCGGGCTGTTT	3325
Db	6213	GGCTTCTGGCGCGACACACGCGCAGGGGGCGTGGCGTTGGCTGGGGCGAGGGCTGTTT	6154
Qy	3326	CATGCCACACCGGTTACTCTTGGCTCGCGCTGGCGCTGAGCGGATGTTGATGAGCG	3385
Db	6153	CACCCCGTCAACCGGATATTCTGCTGCCCTATGCGCGCAGCTGCGGATGCCATCGCGCG	6094
Qy	3386	CTGTTGCCACCGATGCCCTCAGCTCAGCCAAACATATCGAACGCTTTCGCCGTCAGCAG	3445
Db	6093	CGCGACCTGACGACCGCTCGCCCGCTCGCGGGTGGCGGCTGGGCCATCGATCGCGG	6034
Qy	3446	TGGCGCGAAACAGCGATTTTTCGCTGCTGCTAAACCGCATGCTGTTTGGCGGTAAGCGG	3505
Db	6033	GATCGCG---ACCGCTTCTCGCGCTGCTGAACCGGATGCTGTCGCGGCTGCCCGCC	5977
Qy	3506	CAGCAGCGCTGGCGGTGATGCAACGTTTTTACCGGCTCGATGCGCGGTTAATTAGCCGC	3565
Db	5976	GACCGTCTGATCGCTGTCAGCGGTTCTACCGCTTCCGCGCAGCGCTGATCGAGCGC	5917
Qy	3566	TTTTTACCGCGGCACTGCGCTGCGCGATATAACCGCGATTCTGTGCGGCAAGCGCGG	3625
Db	5916	TTCTATGCGGGGCGCTGACATTGGCCGACCGGCTTCGATCGTACCGGACCGCCGCCC	5857
Qy	3626	GTGCCCATCGGTGAAGCGCTGCGCGCTGTTGAATTCGTGGAACAGGGAAGAAAAA	3685
Db	5856	ATTCCGCTGTCAGCGCCGCTGCGCTGCGCTGCCGACGCGCCCTGCTGCGAGGAGAGCA	5797
Qy	3686	TGAAGCGCATTTATGATGAGGCGAGGCTTTGGCGGCTGGCGCTGGCGATTCGCTGCG	3745
Db	5796	TGAGTTTCCGCCATCTGTCATCGCGCAGGTTTCGGCGGGCTTGGCGCTTGCATCCGCTGC	5737
Qy	3746	AGACGCGGGCATACCAACCACTTACTGACGACGCGCACAAACCGGCGGACCGGCT	3805
Db	5736	AATCGCGCGCATTCGCGACCAACCATCGTCGAGCGCGGCGGCGGCGGCGCT	5677
Qy	3806	ATGCTTTGAGGACAGTGGCTTTTACCTTCGATGCGCGAACCCACGGTGATCACCGATCCCA	3865
Db	5676	ATGCTGTGAAACGATCAGGCGCCAGCTCTTCGATGACAGGCGCGGCTGTCGACGACCCG	5617
Qy	3866	GGCCCATCGAAGAGTTGTTCAAGCTGGCAGGAAATTCGCTCAGCTCAGGATTCAGCTCGA	3925
Db	5616	ACAGCTCGGAGAGTGTGGGCCCTCAGCGGCCAACCGATGGAGCGTGACGTGACGCTGC	5557
Qy	3926	TGCCGGTAAACGCTTCTATCGCTGTGTGGAAGATGGCAAAACAGCTTGTATTACGACA	3985
Db	5556	TGCCGGTCTCGCCCTTCTACCGCTGACATGGCGGACGCGCGCAGCTTCGAATACGTGA	5497
Qy	3986	ATAATCAGCGCTGTGAGGACGAGATCGCCACGTTTCAATCCGCAAGATGTAGAAGCT	4045
Db	5496	ACGACGACGACGAGCTGATCCGCCAGTCCCTTCAATCCCGCCGATGTGATGCTG	5437
Qy	4046	ATCGTCAATTTCTTCCTTATTCACGTGAAGTATTTAGAGAGGTTATCTGAAACTCGGCA	4105
Db	5436	ATCGCGCTTCCACGATTACGCCGAGGAGTCTATCGGAGGGGTATCTGAAAGCTGGGA	5377
Qy	4106	CGGTCCGCTTCTGACAGGTGCGTACATGCTCGCGTTCGCGCGCAGTTGGGACGCTGTC	4165
Db	5376	CCACGCTTCTTGAAGCTGGCCAGATGCTGAACGCCGCGCGCGCTGATCGCGCTGC	5317
Qy	4166	AAGCATGCGCAGCGCTCTACAGATGGTGGCGAAATTTATTCAGACGATCATCTGCGTC	4225
Db	5316	AGGCATACCGCTCGGTCCACAGATGGTGGCGCGCTTCATCCAGGACCCGCTCTCGGC	5257
Qy	4226	AGGCTTTTCTTCCACTCATGCTGGTGGCGGTAATCTTTGCAAGCTCATCGATCT	4285
Db	5256	AGGCTTCTGCTTCCACGCTGCTGGTTCGCGGGAAACCCGTTTTCGACAGCTCGATCT	5197
Qy	4286	ATACCTTAATTCATCGCTGAGCGTGAATGGGGGCTGTGTTTCCGCGCGCGCACCG	4345
Db	5196	ATGCGCTGATCCATCGCTGGAAACGCGCGCGCGCTCTGGTTTCGCAAGGCGGACCA	5137
Qy	4346	GGCGCTGTTGAGGCGCATGCGCGCACTGTTTCAGACACTTGGCGGCGAGCTGTACTGA	4405
Db	5136	ACCACTGTTGCGGGCATGTTGCGCTGTTTCAGCGCTTTCGCGCGCACGCTGCTCTGA	5077
Qy	4406	ATGCGGAGTGAAGCCAGCTGGAAACAGCGCAATCGCATTAGCGCGCTTCAGTTAGAG	4465
Db	5076	ATGCGCGCTCAGCGGATTCGACACGAGGGCGATTCGCGCCACGCGCGTCACTGCTGG	5017
Qy	4466	GGGACGACGCTTCGATCGCGCGCTGTGCGCTCCAAATGCCACGCTGTGCTATACG	4525
Db	5016	ACGGCGGCGAGTTGCGCGGGATACGTTGGCCAGCAACGCGCGAGCTGATGCAAGCTATC	4957
Qy	4526	ACAACTGTTGCGCACCATCGCTGGCAATGAACGTCGCGACATGCTGGAAGCGTAAGC	4585
Db	4956	GCGACCTGTGGGCCATACCCGCGCGGCGCACCAAGCGCGCATCTCTGAACCGCGACG	4897
Qy	4586	GATGAGCAACTGCTGTTGTTACTCTATTTGGCGCTGNAATCAGCGCGATGACAGCTCG	4645
Db	4896	GCTGTCGATGTCGCTGTTGCTGCTATTCGCGCTGTCGCGGCGCCGAGAACTGG	4837
Qy	4646	CGCACACACGCTGTTTGGCCGCGTATCTGAGTTGATCGATGAGATTTTCAACA	4705
Db	4836	CCACACAGCGTATCTTCGCGCGCGCTACAGGGGCTGTGACGAGATCTTCAACG	4777
Qy	4706	GCAGCAGCTGCGACAGATTTTTCATTTTACCTGCAACGCGGCTTCGACGAGCGATCCGT	4765
Db	4776	GGCCACGCTGCGGACGATTTCTGATGATCTGATTCGCGCTGCGTGACCGATCCCA	4717
Qy	4766	CGCTGGCACCGCGCTGCGGACGCTTTTATGTTAGCGCGGCTGCGGATCTCGGCA	4825
Db	4716	GCCTGGCGCCGAGGGGATGTCACGCAATACGCTCTTGGCGCGCTTCGCGATCTGGGC	4657
Qy	4826	CGCTGACATCGACTGGCAACAGGAAGGACCGCGCTTGGCGGATCGAATTTTGTCTATC	4885
Db	4656	GGCGCGATGTCGATTGGGAGCCGAGGCGCGGCTATCCCGAGCGCATCTTCGAGGAAC	4597

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2005, 22:47:11 ; Search time 2935 Seconds
(without alignments)
12518.125 Million cell updates/sec

Title: US-10-808-979-18
Perfect score: 6999
Sequence: 1 gcagatggaatgggaatgt.....gacgttttagacatgcaata 6999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 23Sep04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2002bs:.*
8: Geneseqn2003as:.*
9: Geneseqn2003bs:.*
10: Geneseqn2003cs:.*
11: Geneseqn2003ds:.*
12: Geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2623	37.5	6918	2	AAQ06299 Sequence
2	2619.6	37.4	8609	12	Adq14666 Reporter
3	2619.6	37.4	8609	12	Adq26394 Carotenoi
4	2619.6	37.4	8609	12	Adq77262 Nucleotid
5	2619.6	37.4	8609	12	Adq48673 Plasmid p
6	2135.2	30.5	12753	12	Adm98599 Geranylge
7	1900.6	27.2	4624	12	Adp74122 Pantoea a
8	1618.8	23.1	7494	8	Abz69177 Vector co
9	1618.8	23.1	8547	8	Abz69178 Vector co
10	1374.8	19.6	110000	10	Continuation (42 o
11	1374.8	19.6	110000	10	Continuation (7 of
12	1374.8	19.6	110000	10	Continuation (8 of
13	1037	14.8	8625	2	Aav40146 Flavobact
14	1037	14.8	11233	2	Aav40151 DNA seque
15	1035.4	14.8	8625	2	Aat45143 Flavobact
16	827.4	11.8	1493	2	Aax90713 Erwinia u
17	826.2	11.8	1479	2	AAQ06296 Sequence
18	826.2	11.8	1479	2	Aav84081 Carotenoi
19	826.2	11.8	1479	2	Aax19119 Erwinia u
20	826.2	11.8	1479	12	Ado05014 P. anan
21	808.6	11.6	1479	6	Aad35512 Pantoea s

ALIGNMENTS

RESULT 1
AAQ06299
ID AAQ06299 standard; DNA; 6918 BP.

AC AAQ06299;
XX
XX 24-OCT-2003 (revised)
DT 28-JAN-1991 (first entry)
XX
DE Sequence encoding six enzymes of the carotenoid biosynthetic pathway.
XX Carotenoid biosynthesis; vitamin A; cancer; food coloring; ss.
XX Pantoea ananatis.

Key	Location/Qualifiers
mat_peptide	225..1130
mat_peptide	/*tag= a
mat_peptide	/note= "See AAQ06293"
mat_peptide	1142..2435
mat_peptide	/*tag= b
mat_peptide	/note= "See AAQ06294"
mat_peptide	2421..3567
mat_peptide	/*tag= c
mat_peptide	/note= "See AAQ06295"
mat_peptide	3582..5057
mat_peptide	/*tag= d
mat_peptide	/note= "See AAQ06296"
mat_peptide	5096..5983
mat_peptide	/*tag= e
mat_peptide	/note= "See AAQ06297"
mat_peptide	complement(5928..6452)
mat_peptide	/*tag= f
mat_peptide	/note= "See AAQ06298"

EP393690-A.
24-OCT-1990.
20-APR-1990; 90EP-00107493.
21-APR-1989; 89JP-00103078.
05-MAR-1990; 90JP-00053255.
XX

Abt14193 Pantoea s
Acc44762 Pantoea s
Adq14630 Pantoea s
Adq77205 Nucleotid
Adq48640 Pantoea s
Aaq13719 Phytoene-
Aat40792 Phytoene
Aat37094 Phytoene
Aat91544 Erwinia h
Acf70993 Photorhab
Aaq13720 Phytoene
Aat40793 Phytoene
Aat37095 Phytoene
Aat91546 Erwinia h
Adc99016 Bradyrhiz
Aav73181 C. utilis
Aaq13718 Phytoene
Aat40791 Phytoene
Aat37093 E. herbic
Aat41743 Phytoene
Aat91543 Erwinia h
Aat31798 Erythroba
Adg96835 Crticrt2
Adg96934 Crticrt2

Wed Jan 26 08:20:42 2005

PA (KIRI) KIRIN BEER KK.
 XX Misawa N, Kobayashi K, Nakamura K;
 XX WPI; 1990-322212/43.
 XX
 XX DNA sequences encoding enzymes for carotenoid biosynthesis - for prodn.
 PT of carotenoid cpds. e.g. beta-carotene, lycopene, phytoene, etc.
 XX
 XX Disclosure; Fig 7; 4Opp; English.
 XX
 XX Gene products are useful for the synthesis of carotenoids, useful as food
 CC coloring, vitamin A precursor, and possibly in prevention of cancer. See
 CC also AAQ06293-9. (Updated on 24-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 6918 BP; 1522 A; 1876 C; 1881 G; 1639 T; 0 U; 0 Other;

Query Match 37.5%; Score 2623; DB 2; Length 6918;
 Best Local Similarity 64.6%; Pred. No. 0;
 Matches 4006; Conservative 0; Mismatches 2170; Indels 21; Gaps 6;

QY 400 CTTGAACATTACTGCTCCGACAGCAAAACGATCGCGTGGCGCGGATGCGTCC 459
 DB 303 CTTGATCAGTTATTCGCGTGAGGAGAAACGGATGTTGGTGGCGCGATGCGTGAA 362
 QY 460 GGAACTGCGCGAGGCAACGATTCGCTTTTACTGCTGCTGCGAGCGCGAT 519
 DB 363 GGTGCGCTGGCACCAGGAAACGATTCGCCCCATGTTGCTGCTGACCGCCGCGAT 422
 QY 520 ATGGGTGGCAGTGACGCAAAATGGCGTTCTCGATCTCGCTGCGCATGGAATGTTG 579
 DB 423 CTGGGTGGCTCTAGCCATGACGATTTGCGCTGCGTGGAAATGCTC 482
 QY 580 CAGCGGCAATGCTGATTCGATGATGATTCCTCGATGATGAAACGCGAGATGCTGCT 639
 DB 483 CAGCGGCTGCTGATCTTGAAGATGCGCTGATGCGATGCGATGCGATGCGATGCGGCG 542
 QY 640 GGTGCGCTACGCTGATCGCAATTTGTTGAACGTTGCGGATTTGCGCTGCGCTCGCG 699
 DB 543 GGAGCGCTTACCAATTCATCTCATACGAGAGATGTCGCAATCTGCGGCGGTTGCC 602
 QY 700 CTCCTTAGCGCGCATTTGAAGTATGTCACCCGTTGCTGCTGCTGCTGCTGCTGCTGCT 759
 DB 603 TTGCTGATGAAGCTTTGGCGTAAATGCGATGCGATGCGCTGCGCTGCGCTGCGCAAA 662
 QY 760 TCTGAAGCATGCTGAACTCTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 819
 DB 663 AATCGGCGGTTTCTGAACCTGTCAAACGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 722
 QY 820 TTCAGGATCTGACACGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 879
 DB 723 TTTCAAGGATCTGTGAAGGGATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 782
 QY 880 GAACTGAAACCGAGCTGCTGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 939
 DB 783 CACTTTAAACCGAGCGCTGTTGTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 842
 QY 940 GCTTCACGCGAGTGGCGCAAGACTTTAGCTTCTTCCCGCGCGCGCGCGCGCGCGCGCG 999
 DB 843 GCCTTCAGCGAGCGCGTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
 QY 1000 CAATGCTGCAACCTCG 1059
 DB 903 CAATGCTGCAACCTCG 962
 QY 1060 CAGGCGCAATCCAGCTGTTACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1119
 DB 963 GCCGGTAAATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
 QY 1120 GATCACTGCG 1179
 DB 1023 CAACATCTTACGCTTGGCAGTGAGCATCTCTCTGCGCGCTGCGCGCGCGCGCGCGCGCG 1082

QY 1180 CGCAATATATGACGCGCTGTTTAAATCAACAGCTAGCGATATTAACATGAGCGCGCTC 1239
 DB 1083 CAACATTTTATTCAGCGCTGTTTGAACAAAACCTCGCTGCGCTAGTTAAGGATGCTGC 1142
 QY 1240 AGCGGTGGGCGCATTTTGGGTGATCGCGCGCGCTCTACAGCGCACTTTACGCGGTGC 1299
 DB 1143 A-----TGAGCCATTTTGGCGCGATCGCACCGCTTTTACAGCCATGTTTGGCGCATAC 1197
 QY 1300 AGCGGTAGCACAACCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1359
 DB 1198 AGAATCTCGCTCAGAACTGCTCGCGCGCGCTCATCGGTGACCTTTATTCAGCAATACG 1257
 QY 1360 ATGCCCCGCACTTTGCTTAGCAGCAACGATCGAATTTGTTGCTGCTGCGCGCGCGCGCG 1419
 DB 1258 ATATTAACACTTGTATCGATAGCAACCAATGGAATTTCAATCCGTCGGGACAGACGCG 1317
 QY 1420 ATCTGCGCGTTCGCTGCGCGCGCGCTGCTGATCGCTGCGCTGCGCTGCGCGCGCGCTGCG 1479
 DB 1318 ATCCCCCGCGCGCTTAAACGCGCGTGTCTACCTTGGCGCTCATCTCTGCGCGCGCTCAA 1377
 QY 1480 TGTTCGCGTATCGACGATCTCGCGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1539
 DB 1378 TGCTGAAGCTCATCAATGAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
 QY 1540 CGGTACTGAAGCATTTGAACATCGATGCGCGTGTATGCGCGCGCGCGCGCGCGCGCGCG 1599
 DB 1438 AGGCATTTAAGCATCTGCGCGTGTATGCGCGTGTATGCGCGCGCGCGCGCGCGCGCGCG 1497
 QY 1600 GATTGCTGCTGAGCGCTGCTGCTGCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 1659
 DB 1498 CGCTGCTGCTGAGCGCTGCGCGTGTATGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 1557
 QY 1660 ATCGTGAAGCGCGGATTCGCGTTCGCGTGTATGCGCGCGCGCGCGCGCGCGCGCGCGCG 1719
 DB 1558 ATCGTGAACCGGATATGCGCGTGTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1617
 QY 1720 CGCTGAACGCTTTTTCAGCG 1779
 DB 1618 CTCGCGAAGCTTATGCG 1677
 QY 1780 ACGTGATCTCAAAACACG 1839
 DB 1678 GTGTCATTTGCG 1737
 QY 1840 GCCTGTCG 1899
 DB 1738 GTTTTTCG 1797
 QY 1900 AACTGCG 1959
 DB 1798 CGTTACCGCGCTGTTTTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1857
 QY 1960 TCCATGCG 2016
 DB 1858 CTTTCATCCGTTTATTTTATCTCTCAGAAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1917
 QY 2017 TCAAGCGCATGCTTTCG 2076
 DB 1918 TTCAGCG 1977
 QY 2077 TATCGCTGCTGATCG 2136
 DB 1978 GTGAGCTCTGTTAGCG 2037
 QY 2137 TCGCTGCG 2196
 DB 2038 GAAGCG 2097
 QY 2197 AGCTGTTTATCTCATGCG 2256
 DB 2098 AGCTGCGGATTCACCG 2157
 QY 2257 CGATGCTGCG 2316

[illegible]

QY	5557	TGCGTTACTGCTATCACGTCGCGGGCGTGGTTCGGTTTGATGATGGCGCGCGTAAATCGGCG	5611
DB	5451	TGCGCTATTGCTATCACGTTGTCAGGGGTTGTTCGGCTTGATGATGGCGCAAAATCATGGGG	5510
QY	5617	TGCGCAGCAAGCGGTGCTCGATCACGCCCTGCGATTTAGGACATGGCGGTTCCAGCTCACTA	5676
DB	5511	TGCGGGATACGCCACGCTGGAACGGGCCCTGTGACCTTGGGCTGGCATTTCACTTGCACCA	5570
QY	5677	ACATTGGCGCGACATTTGTAGAAAGATGCCAAAATGGTTCGCTGCTATCTGCGCGCAATTCCT	5736
DB	5571	ATAATTGCTCGGATATTGTGGACGATGCGCATGCGGGCGCGCTGTATTCTGCGCGCAAGCT	5630
QY	5737	GGCTCGATCAGGCGGGATTACGCGCGGATACGCTGACTGCACCGCAACATCTGTGCAAGCG	5796
DB	5631	GGCTGGAGCATGAAGGTCTGAAACAAAGAGAAATTATGCGGCACCTGAAAACCGTCAAGCGC	5690
QY	5797	TCGCGCTCACTGGCAGCGCGTTTATGTCGGCGAGCGGAACCTTATTATCACTCGCGCGCAT	5856
DB	5691	TGAGCCGTTACGCCCGTTCGTTTGGTGCAGGAACAGAAACCTTACTATTGTCCTGCCACAG	5750
QY	5857	CCGGTTTACCGGGTTTACCGTTCGGCTCGGCGTGGCCCATCGCTACGGCTCGCGGGCTTT	5916
DB	5751	CCGGCCTGGCAGGGTTGCGCTCGCTTCGCTTCGCGCTGGGCAATCGCTACGGCGAAAGCAGGTT	5810
QY	5917	ATCCGAAATTTGGCGTCAAAAGTTTCAGCACGCCCGGTGTCACGGCTGCGGATTCACGGCAGC	5976
DB	5811	ACCGAAATAAGTTGTCAAGTTGAAACAGGCCGCTCAGCAAGCCTGGGATCAGCGGCAGT	5870
QY	5977	GCACGATPAAAGGTGAAAACTGGCGCTGCTGTGTGTAAGGGGCGAGGTTTGGCGATCACTT	6036
DB	5871	CAACGACCAACCGCGAAAAATTAAACGTGTGCTGTGCGCGCTCTGTGTGTCAGGCGCTTACTT	5930
QY	6037	CGCGTGTCTCGTCTGAAACCGCGTCCGGCTGCTGTGTCGACGCTCTCGTTGATTTT	6096
DB	5931	CCCGATCGGGCTCATCTCTCCCGCCCTGGCGATCTCTGGCAGCGCCG-----TCT	5984
QY	6097	ACGTCCTGTGACGCTGGCGCAGCGTGGCTTGCAGCTTATTACGGGTCGCGCGTAGAGAA	6156
DB	5985	AGCGCCATGTCTTCCCGAGCGTCGCTGAAGTTTTCACAGGGGCGCGCATAGAGAA	6044
QY	6157	ACCAACGACAGCGAGCTTCAACCGCGCGCACCGCATGATGATCGGTGGCGCATGTA	6216
DB	6045	GCCAAAGANAACAACCTTTTGGCCCTGACGGCGTGTATGATACGGTGGCCATATA	6104
QY	6217	TAAGCGCTTAAGATAGCTTTGCGCGGGATATACCGGAACGGCGCAGCGTTGATGCACAG	6276
DB	6105	CBAACGTTTGAGGTAGCCCTTTCGTTGGAAATATACGGAAATGGCCAAAGTTGATGCACAG	6164
QY	6277	GCATCTGTGCAACATGAAGTAGAGCGCGCTAGCTCGTCAATTCGGGCAACCATCCACTG	6336
DB	6165	CCCGTGTGTGCAACATAAATAGAGTAATCCATACGCCGTATACCTGCGCCAAATCCACTG	6224
QY	6337	CAGCGGCCACATGCCCTTCACACCGACATAAATCAGACAAATCGCCAGTACCGCAAAAC	6396
DB	6225	GAGCGGCCACATTCCTGTACTGCCCCAGNTAATCAGCAGGATCGATAATGCAGCAAAAC	6284
QY	6397	CACCGCATAAAGATCGTTGAGCTCAAACTTACCGCTGTGCGGTTTCATGGTGGCAGCATG	6456
DB	6285	CACGCGATAAAGATCGTTAACTTCAACGCCACCTTACCGCGGTTTCATGATGGAAGATG	6344
QY	6457	CCAGCCCCATCCCCAACCGTGCATGATGATTTATGCGACAGCGCGCTACGATTTCCAT	6516
DB	6345	CCATCCCCAACCCAGCGGTGCATGATGATTTGTGTGCCAGTGCAGCAATCACTTCCAT	6404
QY	6517	CACCAACCGTTGCAACAAGATGAAGCACGTTCCATTAACAGAGCATTTGTTCTGCTCAAT	6576
DB	6405	GCCATCAAGGTAAACGAAACGATCAGGGCATTCCAAATCCACAATAAATTTCTCGGT	6446
QY	6577	TGTGGAAGAGGAAGTA	6593
DB	6465	AGAGACGCTGCGCAGCA	6481

RESULT 2.

ADQ14666
ID ADQ14666 standard; DNA; 8609 BP.

AC ADQ14666;

XX 23-SEP-2004 (first entry)

XX Reporter plasmid pPCB15 nucleotide sequence SEQ ID NO:43.

XX carotenoid; carotenoid overproducing microorganism;

KW functional isoprenoid enzymatic biosynthetic pathway; plasmid; circular;

KW gene; ds.

XX Synthetic.

OS WO2004056974-A2.

XX 08-JUL-2004.

XX 19-DEC-2003; 2003WO-US041811.

XX 19-DEC-2002; 2002US-0435612P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Cheng Q, Rouviere PE, Tao L;

XX WPI; 2004-525436/50.

XX Novel carotenoid overproducing microorganism comprising genes encoding functional isoprenoid enzymatic biosynthetic pathway comprising disrupted genes such as dead, mreC, and yfH6, useful for producing carotenoid.

XX Example 3; SEQ ID NO 43; 85pp; English.

XX The present invention describes a carotenoid overproducing microorganism comprising the genes encoding a functional isoprenoid enzymatic biosynthetic pathway comprising a disrupted gene chosen from dead, mreC, and yfH6. Also described is a carotenoid overproducing *Escherichia coli* microorganism (II) comprising an upper isoprenoid enzymatic biosynthetic pathway comprising the genes *dxs*, *dxr*, *ygbP*, *ycbB*, *lytB*, *idi*, *ispA* and *ispB*, a lower isoprenoid enzymatic biosynthetic pathway comprising the genes *crte*, *crti*, and *crty*, mutations chosen from a mutation in the *thrs* gene of the 3159 nucleotide sequence of SEQ ID NO:35 (ADQ14658), mutation in the *rpsA* gene of the 2304 nucleotide sequence of SEQ ID NO:37 (ADQ14660), mutation in the *rpoC* gene of the 5454 nucleotide sequence of SEQ ID NO:38 (ADQ14661), mutation in the *yjE* gene of the 1845 nucleotide sequence of SEQ ID NO:39 (ADQ14662), or mutation in the *rhl* gene of the 2676 nucleotide sequence of SEQ ID NO:41 (ADQ14664), where the genes of the lower isoprenoid enzymatic biosynthetic pathway reside on an autonomously replicating plasmid comprising a replicon chosen from p15A and pMB1. (I) chosen from bacteria, yeasts and filamentous fungi such as *Aspergillus*, *Trichoderma*, *Saccharomyces*, *Pichia*, *Candida*, *Hansenula*, *Salmonella*, *Bacillus*, *Acinetobacter*, *Zymomonas*, *Agrobacterium*, *Erythrobacter*, *Chlorobium*, *Chromatium*, *Flavobacterium*, *Cytophaga*, *Rhodobacter*, *Rhodococcus*, *Streptomyces*, *Brevibacterium*, *Corynebacteria*, *Mycobacterium*, *Deinococcus*, *Paracoccus*, *Escherichia*, *Erwinia*, *Pantoea*, *Pseudomonas*, *Sphingomonas*, *Methylobacter*, *Methylobacter*, *Methylococcus*, *Methylosinus*, *Methylococcobium*, *Methylocystis*, *Alcaligenes*, *Synechocystis*, *Synechococcus*, *Anabaena*, *Thiobacillus*, *Staphylococcus*, *Methanobacterium*, *Klebsiella*, and *Myxococcus*, and (II) are useful for producing carotenoids which involves contacting (I) or (II) with a fermentable carbon substrate, growing the carotenoid overproducing microorganism for a sufficient time to produce carotenoid, and optionally recovering the carotenoid from the carotenoid overproducing microorganism. The present sequence represents the reporter plasmid pPCB15 nucleotide sequence, which is used in an example from the present invention.

SQ Sequence 8609 BP; 1918 A; 2325 C; 2309 G; 2057 T; 0 U; 0 Other;

Query Match 37.4%; Score 2619.6; DB 12; Length 8609;

Best Local Similarity 64.7%; Pred. No. 0;
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;
QY 373 CAGCTGCAGCAAAATTTTACAGCGCATCTTTGAACATTTTACTGCTGCGGACAGCAAAAGC 432
DB |||||
QY 1627 CAGTTGCTGGCTGATATCGATAGCGCCTTGATCAGTTACTGCCGTTTCCAGGTGAGCGG 1686
DB |||||
QY 433 GATCGGTGCGTGGCGGATGCGTGGCGGAACTGCTGGCGCAGGCGCAAAAGTATTCTGCTCT 492
DB |||||
QY 1687 GATTGTGGTGGCGGATGCGTGAAGGACGCTGGCAACCGGCAACGATATTCGTCGG 1746
DB |||||
QY 493 TTATTACTGCTCTGCGAGCGCGATATGGGTGCGAGCTGACGCAAAATTCGCTTCTC 552
DB |||||
QY 1747 ATGCTGCTGTTTAAACAGCGCGGATCTTGCTGTGCGATCAGTCCGCGGATTTACTG 1806
DB |||||
QY 553 GATCTGCGCTGTGAGTGGAAATGTCACGCGGATCGCTGATTTCTGGATGATGATTCCTC 612
DB |||||
QY 1807 GATTTAGCTGCGGTTGAAATGTCATGCTGCTGCTGATTTCTGGATGATATGCCCC 1866
DB |||||
QY 613 TCGATGGAATACGCGCAGATGCTGTCGTCGCTTACCGTCATCGCAATTTGGTGAA 672
DB |||||
QY 1867 TGCATGGAATGTCGCGCAGATGCTGCGGGGCGTCCACCAATTCACACGAGTACGCGTGA 1926
DB |||||
QY 673 AACGTGGGATTTCTCGCGCATCGCGTCTTAGCGCGCATTTTGAAGTGAATGTCATT 732
DB |||||
QY 1927 CATGTGGGATTTCTGCGCGGCTCGCTTACTCAGCAAGCGTTTGGGTGATTTGCGGAG 1986
DB |||||
QY 733 GCACCGCGTTTGCCTGCCATACATAAATCTGAAGCGATTTGCTGAATCTTCGCTGCCGCT 792
DB |||||
QY 1987 GCTGAAGGTCTGACGCGGATGCCAAACTCGCGGGTGTGCGAGCTGTCCACTGCGATT 2046
DB |||||
QY 793 GCGCTGCGAGGCTTGTAGTCAAGGGCAATTCAGAGATCTGCACAGCGGACGCGAGCGCGC 852
DB |||||
QY 2047 GCGATGCGAGGCTGTGCTTTCAGGGCCAGTTTAAAGGACCTCTCGGAAGCGGATAAACCCGC 2106
DB |||||
QY 853 AGCCCGGAGCGATCGCCATGACCAACGAATGAAACAGAGCTGCTGTTTTCGCGCACG 912
DB |||||
QY 2107 AGCGCGATGCCATCTGCTTAACCAATCAGTTTAAACAGCACGCTGTTTTCGCGTCA 2166
DB |||||
QY 913 CTGCAAAATGGCGGATTTGCGCTGACGCTTACCGCAGGTGCGGCAAAAGCTTAGCTTC 972
DB |||||
QY 2167 ACGCAAAATGGCGTCCATTTCGCGCCAAAGCGCTCTCGGAAGCGGTGAGAACCTGCTCGT 2226
DB |||||
QY 973 TTGCGCGCAGGATTTGGCGCAGCGGTTTAACTGCTGCAGCAGCTTCGCCAGCGTTCGAAA 1032
DB |||||
QY 2227 TTCTCGCTCGATCTCGCGCAGCGCTTTCAGTTGCTTGCAGATCTTACCGATGCGCATGACC 2286
DB |||||
QY 1033 CACACGGTAAAGATGTGCACAGGATCAGGCAAAATCCAGCTGTGTACAGATGCTCGGT 1092
DB |||||
QY 2287 GATACCGGAAAGACATCANTCAGGATGAGGTAAATCAACGCTGTGTCATTTATTAGGC 2346
DB |||||
QY 1093 GCTGACGGCGGAAAGCTGCGCTGCGCGATCACCTGCGCAGCGCAGATGCACACCTTGCC 1152
DB |||||
QY 2347 TCAGCGCGGTGCAAGAACGCTGCGACAGCATTTTGCCTGSCCAGTGAACACCTTTCC 2406
DB |||||
QY 1153 TCGCGCTGCGATCGCGGATCGCCACTCGCCAA---TATATGACCGCGCTGTTTAAATCAA 1209
DB |||||
QY 2407 GCGGATGCAAAAGCGGCAATTTCCACACCACTTTTATTTCAGGCGCTGGTTTGAACA 2466
DB |||||
QY 1210 CAGCTAGCGATATTCAACTGAGCGGCTCAGCGGTGGCCACTTTTGGGTGATCGCGC 1269
DB |||||
QY 2467 AAATCGCTGCGCTGAGTTAAGATGCTGCA-----TGAGCCATTTTGGGTGATCGCAC 2521
DB |||||
QY 1270 CGCGCGCTTACAGCACTTTTCAAGCGTTTTCAGCGGTTTTCAGCGGTTTTCAGCAAAAGCTGCTGGCGCG 1329
DB |||||
QY 2522 CGCGCTTTTTCAGCACTGTTTCGCGCTCTCAAAACCTTCTCAGGAATTTAGTGGCGCGC 2581
DB |||||
QY 1330 GGCATGCAATCAATTCATCAGAGCGCGATGCGCGCATCTTGTCTAGGAGCAAGCA 1389
DB |||||
QY 2582 GTCATCGTGTACTGTTTTTTTTCAGCAACATGACTGCAAAAGCGGTGTTAAACGGGCGAGATA 2641
DB |||||
QY 1390 TCGATTTTTCGCGTTCGCGCAACAGACGATCTCTCGGTTTCGCTGGCGCGCTGTTTC 1449
DB |||||

Db 2642 TCGGATTCAGACCGCTCGAGCTGCAAAAGCGATCTCCCGGTTCTTTATCGCACCTGCTGC 2701
Qy 1450 ATCGGCTGGCTCGCGGCGGCTGTGCTGTTTCGGGTGATCAGCATCTCGGTCCT 1509
Db 2702 ACCTGGCGCGCACCACTCGGACCTCGATGTTACGACTGATCAATGAATGGCACGTA 2761
Qy 1510 GCACCGATATGTTGTCGCGCAACTGCTTCGCGTACTGAAAGCAATTGAACATCGATGGCG 1569
Db 2762 CCAGCGATATGTTTTCGCGGAACCTGCCCGCGCTTTTCATGCGTTGAGATAGAGGGG 2821
Qy 1570 TGATCGCGACGAATGGAAGCGGCGGCGGATGCTGCTGAAGCGCTGCAATCTGCCGT 1629
Db 2822 TGATCGTTGATCAAAATGGAGCGGCGAGGTGCAGTAGTCGACAGCGTCAAGTCTGCCGT 2881
Qy 1630 TTGTTTCGGTGGCTGGCTTCGCGTCAATCGGTGAAGCGGGATTCGCTTCGCGTGA 1689
Db 2882 TTGTTTCGGTGGCTGGCGCTGCGCTCAACCGGAACCGGGTTTGCTCTGGCGGTGA 2941
Qy 1690 TGCCCTTCGTTTTCAGAGGATGAAAGCGCTGAAACGTTTTCAGGCGCAGCAGCGATA 1749
Db 2942 TGCCCTTCGAGTACGGCACCGGATCGGCTCGGNAAGCTATACCAACGCGAATAA 3001
Qy 1750 TCTATGATGCAATCATGCTGCTACGGCGACGTCATCTCAAAACAGCGCGGGGCTTTA 1809
Db 3002 TTTATGACTGGCTGATGCGACGTCAGATCGTGTGATCGCGCATCATGCAATGCAATGG 3061
Qy 1810 ATTTGACGAGCGCGCGGATTTACATCAGTGCCTGTCGCGCTGGCAGCAAAATCAGCCAGA 1869
Db 3062 GTTTAGCCCCGCGTGAATACTGCAATCATGTTTCTCCACTGGCAGCAAAATCAGCCAGT 3121
Qy 1870 TGGTGGCGGCTTGTATTTCCACGTCAGCAACTGCGCGCTGCTATCAACGCGCTGGGCG 1929
Db 3122 TGATCCCCGAACTGGATTTTCCCGCAAGCGCTGCCAGCTGCTTTCATCGGTGGAC 3181
Qy 1930 CACTCGCG - - - CCGCGTTTCTCTGCGCGCTGATCGCGCTGCGCGCTGGCGAGCGTGCCTC 1986
Db 3182 CGTTAGCGCAACCCAGGCGGACGCGGGGTCTCAACTCTTTATTTCCGTCGCCGACA 3241
Qy 1987 AGCGGTGTTTATGCTCGCTGGGTACGCTGCGAGCGCATCGCTTCGCGCTGTTCTGCG 2046
Db 3242 AACCCGTAATTTTGCTCGCTGGGACCCCTGGAGGACATCGTTATGSCCTGTTTCAGGA 3301
Qy 2047 ATCTGCGCAGGCGTGCGCCAGCTCGGCTATGCTGTGATGCGCCATGTTGGGGAT 2106
Db 3302 CCATCGCCAAAGCTCGGAAGGTGATGCGCAGTTACTGTTGGCACACTGTGGCGGC 3361
Qy 2107 TAAAGCCGAAACAGAGCATCAGCTGAGCTCGCTGGCGGCGGTGGGTGACGGATTTCG 2166
Db 3362 TCTACGCCACGAGGAGGTGAATGCCCCGGGGGAGCAATTCAGGTTGTGGATTTG 3421
Qy 2167 TCGATCAGCGCGCAGCCCTACAGCACGCGAGCTGTTTATCACTCATGCGGGTTAAACA 2226
Db 3422 CCGATCAATCCGAGCACATTTACAGGCACAGTTGACATCAACATGTTGGGATGAATA 3481
Qy 2227 GCGCGCTGAAGCACATGGAATCGGTPACGCCGATGCTGCGCTGCGGATGCTTTTGTATC 2286
Db 3482 CGGTACTGACGCTATTGCTTCCCGCACACCGCTACTGCGCTGCGCGCTGGCATTTGATC 3541
Qy 2287 AGCCCGCTGGCGCGCGCATGAGTGGCATGACGTTGCTGCGCGCGCATCAGCTTTA 2346
Db 3542 AACCTGGCGTGGCATCACGAATTTGTTTATCATGGCAATCGGCAAGCGTGGCTTCGCTTTA 3601
Qy 2347 GCGCTGTTTCACTAATCGAGCAGCATCTGCAACAGCTGCTGACCGCAGATCGTTACGCG 2406
Db 3602 CTACACGCAATGGCTGGCGCGCAGATTCGATCGCTGCTGACTAACAACCGATTAACCGC 3661
Qy 2407 TAGCGATGTCAGCGATTACGGCGCAGCTGAGCGCGCAGCGGTTCGCAAGCTGCGCGCG 2466
Db 3662 AGCGTATGACAAAAATTCAGGCGCGATTTGCTGTGGCAGCGGCGCACACAGCGCGCGCG 3721
Qy 2467 ACATCGTCAGAGGCGCTGTGCGCAGCAGCAAGTCTGTGCTGGCGGAGGCGACTGATGCG 2526
Db 3722 ATATTGTTGACAGGCGATGCG - GACCTGTACCCAGTACTCAGTGGCGAGGATTTATGCA 3780

Qy 2527 CACGCAATFACGATGTGATTTTGGTTCGGTCTGGACTGCGGAATGCTTGAATTGCGCTGCG 2586
Db 3781 ACCGCACTATGATCTCATTTCTGGTTCGGTTCGGCTATGCTGCTTAATCGCGCTCCG 3840
Qy 2587 TCTCGGTCAATTGAGCGCAACAACTGAATGCTGTTGCTGAGAGCGATGGCATCCGGC 2646
Db 3841 GCTTCAGCAACAGCATCCGATATCGGATCTTCTCTTATTTAGAGCGGGTCTGAGGCGGG 3900
Qy 2647 AGGCAATCATACCTGCTGTTTTCATCAAGCGATCTCAGCGCCGAAACAACTTCGCTGCT 2706
Db 3901 AGGGAACCATACCTGCTGTTTTCATCAAGAGAGATTTAAGCTGAATCAGCATCTGAT 3960
Qy 2707 GCAACCGCTGATTTACCGTTCGTTTATCAGTTCGTTTTCCTGCGCTGCGCGC 2766
Db 3961 AGCGCGCTTCTGTTCCATCACTGCGCGACTACAGGTTCTTCCCGCAACAGCGCTCG 4020
Qy 2767 CAATCGACGCGGATTTATGTTTCATCGCATCAGCGATTTTGGCCGCCATCTTTAGCG 2826
Db 4021 CCATGTGAACAGTGGCTACTACTGCTGACCTCCCGCATTTTCGCGCGGATCTCCCGCA 4080
Qy 2827 GCGGATGGTCAACGATCTGTGCAAAACACAGCGCTTACACAGGTAAACCCACAGCT 2886
Db 4081 ACAGTTTGGACAACTTTATGCTGCATACCGCGTTTCAGCCGTTTCATGCTGAATCGT 4140
Qy 2887 GACGCTGGCGGATGGCGTGAACCTTGTGCGCAAGTGGTGAATGATGTCGCGCTGCA 2946
Db 4141 CCAGTTAGCGGATGGCGGATTTATTCATGCGAGTACAGTGATCGACGACGGGTTACAC 4200
Qy 2947 GCGGAGCGCATCTGAGCTGGGTTATCAGTGTTCCTGGCAAGAGTGGAGCTGCG 3006
Db 4201 GCCTGATTTCTGCACCTACCGGTAGGATTCAGGCAATTTATCGGTACAGGATGGCAACTGAG 4260
Qy 3007 GCGGCGCACGCGCTGACGAGCGCATCTGATGATGCCACCGTTCGATCAGCAAGCGGG 3066
Db 4261 CGCGCGCATGTTTATCGTCAACGATTTATCGATGCGAGCTGATCAGCAAAATG 4320
Qy 3067 TTATCGTTTGTCTACGCTGCGCTCAGCGCGCATCGGCTTATGATTTGAATGATACCA 3126
Db 4321 CTACCGCTTGTGTTATACCTCGCGCTTTCGCGCAACCGCACTGCTGTGATCGAAGACACA 4380
Qy 3127 TTACGTTAACAGCCCGCTGGCGGAGACACACGCTGTCAGCACATCGCGCTATGCTG 3186
Db 4381 CTACATTCACAAAGGCTAATCTTTACGCGGAAAGCGCGCTCAGAACATTCGCGATATGC 4440
Qy 3187 CAATCAGCAAGCTGAGCGCTGAGTACGCTGCTGCTGAAGAGCAGCGCATATTACCGAT 3246
Db 4441 TGCGCGACAGGGTTGCGCGTTACAGACGTTGCTGCGGAAAGAACAGGGTGCATTCGCCAT 4500
Qy 3247 TACCTGAGCGGCAACATCGATCGATTTCTGGCAACAGCAGCGCGCGCAAGCTGCGCGG 3306
Db 4501 TACGTTAACGGCGGATAATCGTCACTTTTGGCAACAGCAAC - - - CGCAAGCCTGTAGCGG 4557
Qy 3307 CCGCGCGCGCGCTGTTTTCATGCCACCGGTTTCTCTTCCGCTGCGCGCTGCGCGCT 3366
Db 4558 ATTAAGCGCGCGGCTGTTTTCATCCCAACCGGCTACTCCCTACCGCTGCGCGCTGCGCT 4617
Qy 3367 AGCGGATTTGTTAGCGCGCTGTTTCCCAACCGATCCCTCAGCTCAGCGCAACATATCGA 3426
Db 4618 GCGCGATCGTCTCAGCGCGCTGATGTTTACCTCTTCTCTTTCACGACGATTCG 4677
Qy 3427 AGCTTTGCGCGTCAAGAGTGGCGGAAACAGCGATTTTTCGCTGCTTAAACCGCATGCT 3486
Db 4678 TCACCTTTGCCAGCAACGTTGGCAACAGGGGTTTTTCCGCAATGCTGAATCGCATGTT 4737
Qy 3487 GTTTTGGCGGTAAAGCGCAGCAGCGCTGGCGGCTGATGCAACGTTTTCGCGCTCGA 3546
Db 4738 GTTTTGGCGGACCGCGCGAGTCACTGCTGGGTGATGCGAGCGTTTCTATGCTTACC 4797
Qy 3547 TCGCGGTTTAAATGACCGCTTTTACCGCGGCACTGCGCGCTGCGCGGATTAACCGCGAT 3606
Db 4798 CGAGGATTTGATGCGCGCTTTTATGCGGGGAAACTCACCGTACCGGACCGATCGGCTACGCAT 4857

QY 3607 TCTGTGCGCAAGCCGCGTCCGCTCCGATCGGTGAAGCGCTCGCGCGCTGTTGAAATCTGT 3666
DB 4858 TCTGAGCGGCAAGCCGCGTCCGCTCCGATTCGCGGCAATTCGAGGCAATTAACGACTCA 4917
QY 3667 CGAACCCAGGGAAGAAAATGAACCGACTTATGTGATTGGCGCAGGCTTTTCGCGCGCTG 3726
DB 4918 TCGTTGAAGAGGACATACTAAGAAACCACTACGTTAAATTTGGTGGCGGCTTTGGTGCCCTG 4977
QY 3727 GCGCTGGGCAATTCGCTCGAAGCGCGGCGCATACCAACCACTTACTCGAGCAGCGCGAC 3786
DB 4978 GCACTGGCAATTCGTTTACAGCGCGCAGGTATTCCTGTTTTGCTGCTTGGCAGCGCGAC 5037
QY 3787 AAACCGGCGGAGCGGCTATGTGTTGAGGACAGTGGCTTTACCTTCGATCCGAGACCC 3846
DB 5038 AAGCCGGGTGGCGGCGCTTATGTTTATCAGAGCAGGCGTTTACTTTTTCGACGCGCCT 5097
QY 3847 ACGGTGATCACCAGTCCAGCGCCATCGAAGAGTTGTTTCACTGCGCAGGAAATCGCTC 3906
DB 5098 ACCGTTATCACCAGTCCAGCGCGATTTGAAGAACTGTTTGGCTTGGCCGCTTAAACAGCTT 5157
QY 3907 AGCGATTACGTGAGCTGATGCGGTAAACCGCTTCTATCGCTGTGCTGGGAAGATGGC 3966
DB 5158 AAGGATTACGTGAGCTGTTGCGGTACGCGCTTTTATCGCTGTGCTGGAGTCCGGC 5217
QY 3967 AAACAGCTTGATTAAGCAATTAACGCGCTGCTGGAGCAGACATCCGACGTTCAAT 4026
DB 5218 AAGGCTTTCAATTAAGATTAACAGCAGGCGGCTTGAAGAGCAGATACAGCAGTTTAA 5277
QY 4027 CCGCAAGATGAGAGCTATCGTCAATTTCTTGGCTTATTCAGCTGAAGTATTAGAG 4086
DB 5278 CCGCGGATGTTGCGGTTTATCGAGGTTCTTGAAGTTTATCGCTGCGCTTATTCATGAG 5337
QY 4087 GGTATCTGAATCTCGGCAAGTGGCTGCTGCTGAGGTCGTGACATGCTGCGCGCTCGG 4146
DB 5338 GCTATCTGAAGCTCGGCACTGTGCTTTTATCGTTCAAGACATGCTTCGGCGCGG 5397
QY 4147 CCGAGTTGGAGGCTGCAAGATGCGCGAGGCTCTACAGATGGTGGCGAAATTTAT 4206
DB 5398 CCCCAGTTGGCAAGCTCGAGGATGCGCGAGCGTTTACAGTAAAGTTGCGCGCTACAT 5457
QY 4207 CAGGACGATCATCTGCTCAGCGCTTTCTTCCACTCATCTGCTGTTGGCGGTAATCCT 4266
DB 5458 GAGGATGACATCTTCGGCAGCGGTTTCTTTTCACTCCTCTTATGTTGGGGGGGATCCG 5517
QY 4267 TTTGCAACGCTCATCTATCTTAAATTCATGCGCTGGAGCGTGAATGGGCGGTGG 4326
DB 5518 TTTGCAACCTCTGCTTATACGCTGATTCACGCGTTAGAACGGGAATGGGCGCTCG 5577
QY 4327 TTTCCGCGGCGGCAACCGCGCGGCTGTTGCGAGGCGATGGCGGACTGTTTCGAGACTTG 4386
DB 5578 TTTCCACGCGGTGGAACCGGCTGCTGTTCAATGGCATGATCAAGCTGTTTCAGGATCTG 5637
QY 4387 GCGCGGAGCTGTTACTGAATCCGNAAGTGGCCAGCTCGGAACCGCGCAATCGCAT 4446
DB 5638 GCGCGGAAGTCTGCTTTAAACCGCGGCTCAGTCAATGGAACCGTTGGGACGAAGATT 5697
QY 4447 AGCGGCGTTTCACTGAGGCGGGAACGACGCTTCGATGCGCGCGCTGTTGGCCTCCAATGCC 4506
DB 5698 CAGGCGCTGCTGGAAGACGCGACACGCTTTGAAACCTGCGCGGTGCGCTCGAACGCT 5757
QY 4507 GAGTGGTGCATCTACGACAAATCTGCTTCGCGCACCATCCGCTGGCAATGAAACGTTGGC 4566
DB 5758 GATGTTGTACATACCTTATCGCGATCTGCTGTCTCAGCATCCCGCAGCCGCTAAGCAGGCG 5817
QY 4567 ACATCGCTCAAGGTTAAGCGCATGAGCAACTCGCTGTTGTACTCTATTTTGGCCTGAT 4626
DB 5818 AAAAATCTGCAATCCAAAGCGTATGATAACTCACTGTTGTACTCTATTTTGGTCTCAAC 5877
QY 4627 CAGCGCATGAACAGCTCGCGCACCAACACCGCTCTGTTTGGCGCGGTTTATCGTGAAGTTG 4686
DB 5878 CATCATCAGATCAACTCGCCCATCATACGCTCTGTTTGGCGCAGGCTACCGTGAACCTG 5937
QY 4687 ATCGATGAGATTTTCAACAGCAGCCAGCTGGGAGACGATTTTTCATCTTACCTGACGCG 4746

DB 5938 ATTACGAAATTTTAAACCATGATGCTGGCTGAGGATTTTTCGCTTTATTACACGCA 5997
QY 4747 CCTCTCAGCAGCGATCCGCTCGCTCGCACCGCCCGCTGGCAGCTTTTATGTGTTAGCG 4806
DB 5998 CTTTGTGTCACGGATCCGCTACTGCGACCGGAAGGTTGGCGCAGCTATTATGTGCTGGCG 6057
QY 4807 CCGGTGCGGCAATCTCGGCAACGCTGACATCGACTGGCAACAGGAAGGACCGCGCTTGGCG 4866
DB 6058 CCGTTCACACTTAAAGGCAACGCGAACTCGACTGGCGGCTAGAAAGGACCCGACTGCGC 6117
QY 4867 GATCGAAATTTTGTCTTATCTGAGCAGCACTACATGCTCCGGGATTAGCTCAGCAATTAGTG 4926
DB 6118 GATCGTATTTTGTACTTCTTGGCAACATTAATCATGCTCCCTGGCTTGGAAAGCCAGTTGGTG 6177
QY 4927 ACACACAGATGTTTACGCGCTTGTATTTTCGCGACACGCTGCATGCCCATCACGCTCG 4986
DB 6178 AGCGACGTAATGTTTACGCGCTTCAATTTTCGCGAGGACTCAATGCTTGGCAGGTTTCG 6237
QY 4987 GCGTTTTCGCTGAGCCGATTTTGAACGAGCGCTGTTCCGCCCGCATAAACCGCAT 5046
DB 6238 GCGTTCGCTGTAACCTATTCTGACCCAGAGCGCTGGTTCCGACCAATAACCGCAT 6297
QY 5047 GCGGATATCAGCAATCTCTATCTGCTGGTGGTACGCTATCCAGGCGCGGCGTCC 5106
DB 6298 AAGCACATGTAATCTTATCTGTTGGCGCAGCACCCATCTCTGGCGCGGCAATTCC 6357
QY 5107 GCGGTGATCGGTTCCGCCAAGCCACCGCAGGCTGATCTCGAGGATCGCGCCGAATGA 5166
DB 6358 GCGGTATGCGCTCGCGNAGCGCAGCGCTTAAATGCTGGAGGACCTGATTTGACGA 6417
QY 5167 ATCGACGCTTTTACTTGAAGTAACGCAACCATGCGGTGGGCTCGAAGAGTTTCG 5226
DB 6418 AT---ACGTCACTTGAATCATGCGTCAAAACCATGCGGTTGGCTCGAAAGCTTTG 6474
QY 5227 CCACCGCGCGCAAGCTGTTGATGACACCGCGCGGACGCTGATGCTGTATGCGT 5286
DB 6475 CCACTGCACTCGACGCTTTTCGACGCGCAAAACCGCTCGACGCTGCTGATGCTTACGCT 6534
QY 5287 GGTGCTGCTCAGCTCGATGATGATTGATGGCAAAACGCTGGCGGAAGCGGCACCGCAGC 5346
DB 6535 GGTGCGCGCACTCGCGACGCTTATGACATCAAAACATGCGGCTTTCATGCGGACGCG 6594
QY 5347 ATGCGCTCGAAGACGCGCAGGCACTATGACGATCTGCAAAATGAAACCGCGCGGCT 5406
DB 6595 CCTCTTCGAGATGCTGAGCAGCGCTGACGAGCTTGAATGAAACCGCTCAGGCT 6654
QY 5407 ACAGCGCGCGCACATGATGAAACGCGGCTTTAGGCGGTTTTCAGGAAGTGGCGATCATTC 5466
DB 6655 ACAGCGGTTGCAAAATGACGAGCGCGCTTTTTCGCGCGTTTTCAGGAGGTTGCGGATGCGCG 6714
QY 5467 ACCAGCTCGCGCAACAACTGGCGTTTGTATCATCTGGAAGGCTTTCGATGATGCAACGCA 5526
DB 6715 ATGATATCGCTCCCGCTTACGCTTGCACCATCTGGAAGGTTTTCGATGATGATGCGCG 6774
QY 5527 ACGAACATTTACGAGCTTCGATGACAGCTCGCTTATCTGCTATCATCGTGGCGGCGTGG 5586
DB 6775 AAACGCGCTTACCTGACACTGAGCAGATACGCTGCTTATGCTATCATCGTGGCGGTTG 6834
QY 5587 TCGGTTTGTATGATGCGCGCTGTAATGGGCTGGCGACGAAGGCTGCTGATCAGCTCCCT 5646
DB 6835 TGGGCTGATGATGCGCAAAATTTATGGGCTTTCGCGATTAACGCCACGCTCGATTCGCGCT 6894
QY 5647 GCGATTTAGGACTGCGGTTCCAGCTCACTTAAATTCGCGCGGACATTTGTAGAAAGATGCGG 5706
DB 6895 GCGATCTCGGCTGCTTTCAGTTGACCAACATTTGCGGCTGATATTGTCGAGATGCTC 6954
QY 5707 AAAATGGTGGCTTATCTGCGCAATCTGCTCGATCAGCGGATTAACGCGCGATA 5766
DB 6955 AGGTGGCGGCTGTTTATCTGCTGAAAGCTGCTGGAAGGAAAGGACTGACGAAAGCGA 7014
QY 5767 CCGTGACTGACCGCAACATCTGTCAGCGCTCGCTCCTGCTGAGCGGCTTTAGTGGCGG 5826

Db 7015 ATTATGCTGCGCCAGAAACCGGAGCGCTTAAGCGGTATCGCGGGGAGCTGGTACGGG 7074
Qy 5827 AGCGGGAACCTTATTATCACTCGGCGCGATCCGGTTTACCGGGTTTACCGTGGCTCGG 5886
Db 7075 AAGCGGAACCTTATTACGTATCATCAATGGCGGGTCTGGCACAATTAACCTTACCGTGG 7134
Qy 5887 CGTGGGCATCGTACGGCTCGCGGGTTTATCGGAAATTGGCGTCAAGTTGACGACG 5946
Db 7135 CCGTGGCCATCGCAGCAGGAAGAGGTGTACCGTAAATTTGGCGTGAAGTTGAACAGG 7194
Qy 5947 CCGGTGTGACCGCTGGGATTCACGCGAGCGCAGCAGCAAGTAAGAGTGAAGAACTGGCGTGC 6006
Db 7195 CCGGTAAAGCAGCGCTGGGATCATCGCCAGTCCACGTCCACCGCCGAAATTAAGCGTTT 7254
Qy 6007 TGGTGAAGGGGAGGTTTGGGATCATCTCGCGTGTCTCGTCTCTGAAACCGGCTCGG 6066
Db 7255 TGCTGACGGCATCCGGTACGGAGGTACTTCCCGGATGAAGACGTATCCACCCCGCTCGT 7314
Qy 6067 CTGGTCTGGCAGGCTCTCGTTGATTTTACGTCCGTGACGCTGGCGAGCGTGGCTTG 6126
Db 7315 CTCATCTCTGGCAGCGCC-----GATCTAGCGGATGCTTTCTCTCAGCGTGGCTG 7368
Qy 6127 CAGCTTATTACGGGTGGCGGTAGAGGAAACCAACGACACGCGAGCTTTACGCCCGG 6186
Db 7369 AAGTTTATAGATAACGGTGGCGGTACAGAAACCAAGGACACGCGAGCTCTTTCCCGT 7428
Qy 6187 CACCGATGATGATCGGTGGCGGATGTAAGCGCTTAAGATAGCCTTTGCGCGGAT 6246
Db 7429 TACAGCATGATGATACGGTGGCGGATGTAACCGTTTACGGTACGCTTTGCGCGGAT 7488
Qy 6247 ATAGCGGAACGGCGGCTTGATGACACGAGGCATCGTGCACCATCAAGTAGAGCGCGC 6306
Db 7489 GTAGCGGAACGGCGGCTGGTGTACCACTCGTGGTGGACCAATAAATACAGTAAAC 7548
Qy 6307 GTACGTGCTATTCCGGCACCAATCCATCTGACGCGGCCATCGCTTGACACCGACATA 6366
Db 7549 ATAAGCGGTATGCTGTCACCAATCCATCTGAGCGGCCAGATTCTGTACTGCGGAAGTA 7608
Qy 6367 AATCAGCAATCGCGAGTACCGCAACACCGCATAAAGATGTTGAGCTCAAACTT 6426
Db 7609 AATCAGGGCAATCGACATGGCGAATACCGCATAGATCGTTAACTTCAATGC 7668
Qy 6427 ACCGCTGTGGGTTTATGTTGACAGATGCCAGCCCCCATCCCAACCGGTGATGATGA 6486
Db 7669 GCCTTTACGCGGTTTATGATGTGAAGATGCCAGCCCCCAACCCAGCGTGCATGATGA 7728
Qy 6487 TTTATGGACAGCGCGCTACGATTTCCATCACCACCGGTTGCCAACAAGATAGCAC 6546
Db 7729 TTTATGTGCCAGTGCAGCAACCACTTCCATGCGCACCGGTGACAAACACGATCAGGCG 7788
Qy 6547 GTTCCATAACGAGCATTTGTTGTCGA 6574
Db 7789 ATTCCAAATCCACAAATTTCTCAA 7816

RESULT 3

ADQ26394

ID ADQ26394 standard; DNA; 8609 BP.

XX

AC ADQ26394;

XX

DT 23-SEP-2004 (first entry)

XX

DE Carotenoid reporter plasmid pPCB15.

XX

KW Chromosomal engineering; carotenoid; pPCB15; ds.

XX

OS Unidentified.

XX

PN W02004056973-A2.

XX

PD 08-JUL-2004.

XX

PF 19-DEC-2003; 2003WO-US041810.
XX
PR 19-DEC-2002; 2002US-0434602P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Suh W;
XX
DR WPI; 2004-507710/48.
XX
PT Directed integration of an expressible DNA fragment lacking a selectable
PT marker into a bacterial chromosome comprises co-transforming
PT recombination proficient host with at least two linear recombination
PT elements.
XX
PS Disclosure; SEQ ID NO 37; 96pp; English.
XX
CC The invention provides a method for the integration of foreign genetic
CC elements into a bacterial chromosome without the need of a cloning step.
CC The method relies on the presence of the bacteriophage lambda-Red
CC recombinase system in a recombination proficient bacterial host. At least
CC 2 linear recombination elements or constructs are co-transformed into the
CC host, and are assembled in the correct orientation in the bacterial
CC chromosome by homologous recombination. Selectable markers used for the
CC selection of transformants are later excised by the action of a site-
CC specific recombinase. The method allows for multiple chromosomal
CC modifications within Escherichia coli, which is essential when
CC engineering biosynthetic pathways for industrial purposes. The method is
CC illustrated using isoprenoid/carotenoid biosynthesis as an example. The
CC promoters of key genes that encode the rate-limiting enzymes involved in
CC the isoprenoid pathway are engineered, resulting in increased beta-
CC carotene production. The present sequence is that of carotenoid reporter
CC plasmid pPCB15, which was used in examples from the invention.
XX
SQ Sequence 8609 BP; 1918 A; 2325 C; 2309 G; 2057 T; 0 U; 0 Other;

Query Match 37.4%; Score 2619.6; DB 12; Length 8609;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;
Qy 373 CAGCTGCAGCAAAATTTTACAGCGCATCTTCAACATTTTACTGCTGCGGACAGCAAGC 432
Db 1627 CAGTTGCTGGCTGATATCGATAGCGGCTTGTACTCAGTTACTGCGGTTTCAGGGTGAGCGG 1686
Qy 433 GATCGCGTGGTGGCGGATCGGTCCGGAACGCTGGCGCAGGGCAACGATTTCGTCT 492
Db 1687 GATTGTGGTGGTGGCGGATCGTGAAGGACGCTGGCAGCGGCAACGATTTCGTCCG 1746
Qy 493 TTATTACTGCTGCGCAGCGCGGATGATGGTTCGAGCTGACGCAAAATGGCGTTCTC 552
Db 1747 ATGCTGCTGTTATTAAACAGCGCGGATGTTGGTGTGCGATCAGTCAAGGGGATCTG 1806
Qy 553 GATCTCGCTGTGCAAGTGAATGTGACGCGGATCGCTGATTTCGTGATGACATTC 612
Db 1807 GATTAGCTGCGCGGTTGAAATGTTGATGCTGCTCGCTGATTCTGATGATGCTCC 1866
Qy 613 TCGATGATACGCGCAGATGCTGCTGGTGGCTTACGTCATCGTCGGAATTTGGTGA 672
Db 1867 TGCATGAGGATGCGCAGATGCTGCGGGGGTCCCAATTCACGCGAGTACGGTGA 1926
Qy 673 AACGTGGGATTTCTCGCGCATCGCTGCTTACCGCGCATTTTGAAGTGAATGCCATT 732
Db 1927 CATGTGGGATTTCTGCGCGGCTGCTTACTCAGCAAGCGTTTGGGTGATTGCCGAG 1986
Qy 733 GCACCGGTTTCCCTGCCATACATAAATCTGAAGGATGTGTGAATCTCCGCTGCCGTC 792
Db 1987 GCTGAAGGTCTGACGCGGATAGCCAAAATCTCGCGGGTGTGCGAGCTGTCCATCGATT 2046
Qy 793 GGCCTGCGAGGCTTAGTGCAAGGCAATTCAGGATCTGCAGCGGACGCGACGCGCCG 852
Db 2047 GGCATGCGAGGCTGTGTTTCAGGGCAGTTTAAAGACCTCTCGAAGGCGATAAACCCCG 2106
Qy 853 AGCCCGGAAGCGATGCCATGACCAACGAACTGAAACACGAGCTGTGTTTCGCGCCACG 912

Db 2107 AGCGCGATGCCATACTACTGTAACCAATCAGTTTAAACACGACCGCTGTTTTGCGCGTCA 2166
QY 913 CTGCAAAATGGCGCGGATGCGCGTTCAGCGCTTCACGCGAGGTGCGGCAAGACTTAGCTTC 972
Db 2167 ACGCAAAATGGCGTCCATTGCGCCCAAGCGTCTCGAAGCGGTGAGAACCTGCATCGT 2226
QY 973 TTGCGCCAGGATTTGGCCAGCGGTTTAACTGCTGAGAGACTCGCGACGCGTTGCAAA 1032
Db 2227 TTCTCGCTCGATCTCGGCCAGCGCTTTCAGTTGCTTGACGATCTTTACCGATGCGATGACC 2286
QY 1033 CACACGGTAAAGATGTGCACAGGATCAGGCAAAATCCACGCTGCTGATCAGATGCTCGGT 1092
Db 2287 GATACCGGCAAGACATCAATCAGGATGAGTAAATCAACGCTGCTCAATTTATTAGGC 2346
QY 1093 GCTGACGGCGGAAAGCTGCGCTGCGGATCACTGCGAGCGCAGATGCACACCTTGCC 1152
Db 2347 TCAGGCGGTGCAAGAACGCTGCGACAGCATTTGGCGCTGCGCAGTGAACACCTTCC 2406
QY 1153 TCGCGCTGCGATCGCGGATCGCACTCGCAAA---TATATGACGCGCTGTTTAAATCAA 1209
Db 2407 GCGGATGCAAAAGCGGCAATTCACACCCAACTTTTATTCAGGCGCTGTTTGACAAA 2466
QY 1210 CAGCTAGCGATATTCAACTGAGCGGCTCAGCGGTGCGCACTTTGCGGTGATCGCG 1269
Db 2467 AAATCGCTGCGCTCAGTTAAGATGCTGCA-----TGAGCCATTTTGGCGGTGATCGCAC 2521
QY 1270 CGCGCGCTTACAGCCACTTTACGCGTTTGCAGCGGTTAGCAAAACGCTGCTGGCGCGG 1329
Db 2522 CGCCCTTTTACGCCATGTTGCGGCTCTGCAAAACCTTGCTCAGGAAATAGTGGCGCG 2581
QY 1330 GCATCGCATCAATTCATCAGCAAGCGGATGCGCGCACTTTGCTTAGCGACGAACGCA 1389
Db 2582 GTCATCGTTTACGTTTTTTCAGCAACATGACTGCAAAAGCGCTGTGTAACGGCAGCGATA 2641
QY 1390 TCGATTTTGTGCGTTCGCGCAACAGACGATCTCGCGTTTGGTGGCGCGCTGTCG 1449
Db 2642 TCGGATTCAGACCGTTCGAGATGCAAAACGATCTCCCGGTTCTTATCGCACCTGCTGC 2701
QY 1450 ATCGGCTGCGCTCGCGCGCGCTGCTGCTGTTTGGCGTATGCAACGATCTCGCGTCT 1509
Db 2702 ACTTGGCGCGCACCACTCGACCTCGATGTTAGCACTGATCAATGAATGGCACGTA 2761
QY 1510 GACCGATATGCTGTGCGCGAACTGCTGCGGTACTGAAGCATGAACATCGATGGCG 1569
Db 2762 CCAGCGATATGCTTTGCGCGAACTGCGCGCGCTTTTCATGCTTGCAGATAGAGGGCG 2821
QY 1570 TGAATCGCGAAGATGGAAGCGCGCGGATGCTGCTGAAGCTGCACTGCGGT 1629
Db 2822 TGATCGTTGATCAAAATGGAGCGCGAGGTGCGAGTAGTCGCAAGCGCTCAGGTCTGCGGT 2881
QY 1630 TTGTTTTCGGTGGCTGCGCTTTGCGGTTCAATCGTGAAGCGGATTTCCGCTTGGGTGA 1689
Db 2882 TTGTTTTCGGTGGCTGCGGCTGCGGCTCAACGCGAACCGGTTTTCCTCTGCGGTGA 2941
QY 1690 TCGCTTTCGGTTTGCACAGGATGAACAAAGCGCTGAACGTTTTCAGGCCAGCAGGATA 1749
Db 2942 TGCCTTTGAGTACGGCACCGCATGCGGCTCGGGAACGCTATACCAACAGCGAAAAA 3001
QY 1750 TCTATGATCGCATCATGCGTTCGTCAGCGGATGATCTCAACACGCGCGCGGCTTTA 1809
Db 3002 TTATGACTGGTGTATGCGACGTCACGATCGTGTGATCGCGCATGATGATGCAATGG 3061
QY 1810 ATTTGACGAGCGCGCGGATTTACATCAGTGTGCTGCGCGCTGGCAACAAATCAGCCAGA 1869
Db 3062 GTTTAGCCCGGTGAAAACTGCAATCATTTTCTCACTGGCAAAATCAGCCAGT 3121
QY 1870 TGTGCGCGGCTTTGATTTTCAAGTCAAGAACTGCGCGCTGCTATCAACGCGTGGGCG 1929
Db 3122 TGATCCCGAACTGGATTTTCCCGCAAGCGCTGCGACACTGCTTTTATGCGGTGGAC 3181
QY 1930 CACTCGCG---CCCGGTTTCTCTGCGCGCTCCATCGCGCTGCGCGCTGCGCTG 1986

Db 3182 CGTTAGGCAACCCAGGGACGCGCGGTCATCACTTCTTATTTTCGTCGCCGACA 3241
QY 1987 AGCGCGTGGTTATGCTCGTGGGTACGCTGCAAGGCAATCGCTTCGCGCTGTTCTGC 2046
Db 3242 AACCCGCTATTTTGGCTCGCTGGGCACTGTCAGGACATCGTTATGCGCTGTTGAGGA 3301
QY 2047 ATCTGGCGAGCGGTGCGCGAGCTGCGGTATCGCTGGTATCGCCATTCGTTGGGGAT 2106
Db 3302 CCATCGCCAAAGCTTGCAGAGGTGGATGCGAGTTACTGTTGGCACATGTCGCGGCG 3361
QY 2107 TAAACCGCGAACAGACGATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGATTTGCG 2166
Db 3362 TCTAGCCACAGCAGCAGGTGACTGCGCGCGGCGGACATTCAGGTTGTGATTTTG 3421
QY 2167 TCGATCAGCGCGAGCCCTTACAGCAGCGCAGCTGTTTATCTACTCATGCGCGGTTAAACA 2226
Db 3422 CGATCAATCCGACGACTTTTACAGGACAGTTGACAATCACATGTCGCGATGAATA 3481
QY 2227 GCGCGCTGGAAGCACTGGAATGCGGTACGCGATGCTGCGCTGCGGATTTGATGTC 2286
Db 3482 CGGTACTGAGCGTATGCTTCCGCAACCGCTACTGCGCTGCGCTGGCATTTGATC 3541
QY 2287 AGCCCGGCTGGCGCGCATTTGAGTGGCATGAGTTGGTGGCGCGCATCACGCTTTA 2346
Db 3542 AACCTGGCGTGGCATCGAATTTTATCATGGCATCGGCAAGCTGCTCTCGTTTA 3601
QY 2347 GCGGTGTTTCACTCACTGAGCAGCATCTGCAACAGCTGCTGACCGAAGCTTTAGCGCG 2406
Db 3602 CTACAGCCATGCGCTGGCGGCGAGTTGATGCTGCTGCTAACAACCGATTACCCGC 3661
QY 2407 TACGATGTACGAGTTTACGCGGAGCTGACGCGGCGAGCGGTGCGAGCTGCGCGCG 2466
Db 3662 AGCGTATGACAAAAATTGAGCGCGCATTTGCGTCTGCGAGCGGCGCACACGAGCGCGCG 3721
QY 2467 ACATCGTCGAGCAGCGCTGTGSCAGCAGCAAGTGTGCTGGCGAGGCGACTGATGCG 2526
Db 3722 ATATGTTGAAACAGCGATGCG-GACCTGTGAGCCAGTACTCAGTGGGAGGATATGCA 3780
QY 2527 CACGCAATACGATGTGATTTTGGTGGTGTGACTGCGAATGCTGATTTGCGCTGCG 2586
Db 3781 ACCGCACTATGATCTCATTTGGTGGTGGCTGCGCTGCTGCTAATGCTTATCGCTCGC 3840
QY 2587 TCTGGTCAATGACGCAACAACTGAAATGCTGTTGTTGGAGAGGCGATGCGCATCCGC 2646
Db 3841 GCTTCAGCAACAGCATCCGGATATCGGATCTTGTCTTATGAGCGGGGCTCTGAGCGCG 3900
QY 2647 AGGCAATCATACTCGTGTGTTTCAACAGCGATCTCAGCGCGCAACAACTTCGCTGGCT 2706
Db 3901 AGGGAACCATACCTGCTCTTTCAGAGAGGATTTAAGCTGAATCAGCATCGCTGGAT 3960
QY 2707 GCAACCGCTGATTACCGTGGTGGTTCAGGTTTATCAGTGGCGTTTCTGCGCTGCGCGC 2766
Db 3961 AGCGCGCTTGTGGTCCATCACTGCGCGACTACCAAGTTTCTTCCCGCAACGCGCTCG 4020
QY 2767 CAATCTGACGCGGATTTATGTTCCATCGCATCAGCGGATTTTGGCGCGCATTTTACGC 2826
Db 4021 CCATGTGAACAGTGGCTACTACTGCTGACCTTCCGCGCATTTTCCGCGGATATCTCGGCA 4080
QY 2827 GSCGATGGTGAAGTCTGTGCAAAACACAGCGCTGACAAAGGTAAACCCACGAGT 2886
Db 4081 ACAGTTTGAACAACTTATGCTGATACCGCGTTTTCAGCGGTTTCTGCTGAATCGGT 4140
QY 2887 GACGCTGGCGATGGCGGTGAACCTTGTGCGCAAGTGGTGAATGATGGTGGCGGCTGCA 2946
Db 4141 CCAGTTAGCGGATGGCGGATTTATTCATGCGCAGTACAGTATCGACGCGGGTTACAC 4200
QY 2947 GCGGACGCAATCTGAGCTGGGTTATCAGGTGTTTCTTGGCAAGAGTGGCAGTGGC 3006
Db 4201 GCTGATTTCTGCACTACGCGTAGGATTCAGGCAATTTATCGGTGAGGAGTGGCACTGAG 4260
QY 3007 GCAGCGGCAACGCGCTGACAGCGGATCTCTGATGATGCGCACCGCTGATCAGCAGCGG 3066
Db 4261 CGCGCGCATGTTTATCGTCCCGATTTATGATGCGAGCGGTGATCAGCAAAATGG 4320

QY 3067 TTATCGTTTGTCTACACGCTGCGCTCAGCGCCGATCGGCTATTGATTGAAGATACCCA 3126
DB 4321 CTACCGCTTTGTTTATACCTGCGCTTTCCGCAACCGGACTGCTGATCGAAGACACACA 4380
QY 3127 TTACGTTAAACAGCGCGGCTGGCGGAGAACACCGCTCGTCAGCACATCGCGGACTATGC 3186
DB 4381 CTACATTGACAGGCTAATCTTCAGGCCGAACGGGGCGTCAGAACATTCGGGATTATGC 4440
QY 3187 CAATCAGCAAGGCTGGAACGCTGAGTACGCTGCTGCTGAGAGACGCGCATATTACCGAT 3246
DB 4441 TCGCGCACAGAGGTTGGCGTTTACAGACGTTGCTCGGGAAGAACAGGGGTGCAATCCCAT 4500
QY 3247 TACCTTGACGCGAACATCGATCTGTCGCAACAGCAGCGCGCAAGCGTGCAGCGG 3306
DB 4501 TACGTTAAACGGGCGAATATCGTCAGTTTGGCAACAGCAAC---CGCAAGCCTGTAGCGG 4557
QY 3307 CCTGCGCGCGGCTGTTTCATGCCACACCGGTTACTCTTTGCGCTCGCGTGGCGCT 3366
DB 4558 ATTAACGCGCGGCTGTTTCATCCGACAAACCGGCTACTCCCTACCGCTCGCGGTGCGCT 4617
QY 3367 AGCGGATGTTAGCAGCGCTGTTGCGCACCGATGCCCTCAGCTCAGCCAAACATATCGA 3426
DB 4618 GCGCGATCTCTCAGCGCGCTGGATGTGTTTACCTCTTCTCTGTTCACACGATTCG 4677
QY 3427 AGCTTTGCGCTCAGCAGTGGCGGAACAGCGATTTTTCCGCTGCTGCTAAACCGCATGCT 3486
DB 4678 TCACITTTGCCAGCAACGTTGGCAGCAACAGGGGTTTTTCCGATGCTGATCGCATGTT 4737
QY 3487 GTTTTTGGCGGTAAACCGCAGCAGCGCTGGGCGGTGATGCAACGTTTTTACCGCTCGA 3546
DB 4738 GTTTTTAGCGGACCGCGCGAGTCAGCTGGCTGTGATGACAGCGTTTTATGGCTTACC 4797
QY 3547 TGC CGGTTAAATTAGCGCTTTTACCGGGCAACTGCGCCCTGCGGGATAAACCGGAT 3606
DB 4798 CGAGGATTTGATTGCGCGCTTTTATGCGGGAACACTCACCGTGACCGATCGGCTACGCA 4857
QY 3607 TCTGTGCGCAAGCGCGGTGCCATCGGTGAAGCGCTGCGCGCTGTTGAAATCTGCT 3666
DB 4858 TCTGAGCGCAAGCGCGCTTCCGCTTCCGCTTTTCGGGATTTGAGGCAATTAAGCACTCA 4917
QY 3667 CGAACAGGGAAGAAAAATGAACCACTTATGTGATTGGCGCAGGCTTTGGCGGCTG 3726
DB 4918 TCGTTGAGAGGACTACATGAACCACTACGTTAATTTGGTGGCGCTTTGGTGGCTG 4977
QY 3727 GCGCTGCGATTCGCTGCAAGCGCGGATACCAACCTTACTCGCAGCAGCGGAC 3786
DB 4978 GCACTGGCAATTGGTTACAGGCGGCAAGTATTCCTGTTTGTGCTTGAGCAGCGCGAC 5037
QY 3787 AAACCGGCGGACCGCGCTATGTTTGAGGACAGTGGCTTTTACCTTCGATGCGCGACCC 3846
DB 5038 AAGCGGTTGGCGGCTTATGTTTATCAGGACGCGGCTTTACTTTGATGACGCGCT 5097
QY 3847 ACGGTATCACCGATCCCAGCGCCATCGAAGAGTGTTCACGCTGGCAGGAAATCGCTC 3906
DB 5098 ACGGTTATCACCGATCCCAGCGGATTTGAAGAACTGTTGCTCTGGCGGTAACAGCTT 5157
QY 3907 AGCGATTACGTCAGCTGATGCGGTAAAGCCCTTCTATCGCTGTGCTGGGAGATGCG 3966
DB 5158 AAGGATTACGTCAGCTGTTGCGGCTCAGCGCTTTTATTCGCTGTGCTGGGAGTCCGCG 5217
QY 3967 AAACAGCTTGAATACCAATATACCGCGCTCTCGAGCAGCAGATCGCCAGTTCAT 4026
DB 5218 AAGGCTTCAATTACGATAACGACAGGCGCCAGTTAGAACGCGCAGATACAGCAGTTTAA 5277
QY 4027 CCGCAAGATGTAGAGGCTATCGTCAATTTCTTTGCGCTATTCAGCTGAAGTATTAGAGAG 4086
DB 5278 CCGCGGATGTTGCGGTTATCGAGGCTTCTTTGACTATTCGCGTGGCTTTCAATGAG 5337
QY 4087 GGTATCTGAAACTCGSCAGGTCGCTTTCTGACGTGCTGACATGCTGCGCGTCGCG 4146
DB 5338 GGCTATCTGAAGCTCGCACTGTGCTTTTATCGTTTCAAGACATGCTTCGGGCGCGG 5397

QY 4147 CCGCAGTTGGACGCTCTGCAAGCATGGCGCAGCGCTCTACAGCATGGTGGCGAAATTTATT 4206
DB 5398 CCCAGTTGGCAAGCTCGAGGCATGGCGAGCGTTTACAGTAAAGTTGCGCGCTACATT 5457
QY 4207 CAGACGATCATCTGCGTCAAGCGGTTTTCTTCCATCATTTGCTGGTGGCGGTAATCCT 4266
DB 5458 GAGGATGAGCATCTTCGCGAGCGGTTTTCTTTTCACTCGCTTCTAGTGGGGGGAATCCG 5517
QY 4267 TTTGCAAGCTCATCGATCTATACCTTAATTCATCGCTGGAGCGTGAATGGGCGTGTGG 4326
DB 5518 TTTGCAACCTCGTCAATTTATACGCTGATTCACGCTTAGAACGGGATGGGCGTCTGG 5577
QY 4327 TTTCCGCGCGCGGACCGCGCGCTGCTGTCAGGGCATGGCGCACTGTTTCGAGGACTTG 4386
DB 5578 TTTCCACGCGGTGGAACCGGTCGCTGCTCAATGGCATGATCAAGCTGTTTTCAGGATCTG 5637
QY 4387 GCGCGGAGCTGTTACTGTAATGCCGAAGTGAAGCAGCTGGAACACGAGCGCAATCGCAT 4446
DB 5638 GCGCGGGAAGTCTGCTTAAACGCGCGGTCACTATATGGAACCGGTTGGGGACAAGATT 5697
QY 4447 AGCGCGCTTCAAGTTAGAGGGCGGACGCTTCGATGCGCGCTGTGGCCTCCAATGCC 4506
DB 5698 CAGCGGTGCAAGTTGGAAGACGCGCAGACGCTTGAACCTCGCGGTGGCGTCGACGCT 5757
QY 4507 GACGTGTGCTACTACGACAAACTGCTTCGCGCAACCTCGCTGCGCAATGAACGTCGCG 4566
DB 5758 GATGTTGTACATACCTATCGCGATCTGCTGTCTCAGCATCCGCGAGCGCTAAGCAGCGG 5817
QY 4567 ACATCGCTGAAGCTAAGCGCATGAGCACTGCTGTTTGTACTCTATTTTGGGCTGAAT 4626
DB 5818 AAAAACTGCAATCCAGCGTATGAGTAACCTCACTGTTTGTACTCTATTTTGGTCTCAAC 5877
QY 4627 CAGCGCATGAACAGCTCGCGCACCAACCTGCTGTTTGGCGCGCTTATCGTGAGTTG 4686
DB 5878 CATCATCAGCATCACTGCGCCATCATACGCTCTGTTTGGGCCAGCTACCGTGAACG 5937
QY 4687 ATCGATGAGATTTTCAACAGCAGCAGCTGGGAGACGATTTTCACTTCACTGACGCGG 4746
DB 5938 ATTACGAAAATTTTAAACCATGATGCTGCTGAGGATTTTTCGCTTATTTTACACGCA 5997
QY 4747 CCTGACAGCGGATCCGCTGCTGCGACCGCGCTGCGCAGCTTTATGTTGTAGCG 4806
DB 5998 CTTGTGTGCGGATTCGCTCACTGCGACCGGAGGTCGCGAGCTATTTATGCTGCGG 6057
QY 4807 CCGGTGCGCGCATCTCGGCACCGCTGACATCGATGGCAACAGGAAGGACCGCGCTGCGC 4866
DB 6058 CTTGTTCCACTTAGSCACGCGGAACCTCGACTGGGCGGTAGAAGGACCCGACTGCGC 6117
QY 4867 GATCGAATTTTGTCTTATCTGAGCAGCAGCTACATCGCGGATTTACGTCAGCAATTAGT 4926
DB 6118 GATCGTATTTTGTACTACTTTGAGCAACATTTACATGCTGCTGCTGCGAAGCCAGTTGGT 6177
QY 4927 ACACACAGATGTTTACCGCTTTGATTTTTCGCGACAGCTGATGCTGCGCATCACGGCTCG 4986
DB 6178 ACGCACCGTATGTTTACCGCTTTGATTTTTCGCGACAGCTCAATGCTTGGCAAGGTTGCG 6237
QY 4987 GCGTTTTGCTGAGCGGATTTTGAACGAAAGCGCTGCTTCCGCCCGCATAACCGCAT 5046
DB 6238 GCTTCTCGTTGAACCTTATCTGACCGAGCGCTTGGTTCGACACACATAAACCGCAT 6297
QY 5047 GCCGATATCAGCAATCTCTATCTGTTGGTGGCGGTACGCAATCCAGCGCGGCGCTGCC 5106
DB 6298 AAGCACATTTGATAATCTTTTATCTGTTGGCGCAGCACCCATCTTGGCGCGGCAATCCC 6357
QY 5107 GCGGTGATGCGTTTCGCGCAAGGCCACCGCAGGCTGATGCTGAGGATCGCGCCGAATGA 5166
DB 6358 GCGGTAATCGGCTCGCGAGGCGAGCGGCTTAATGCTGGAGGACCTGATTTGACGA 6417
QY 5167 ATCGACAGCTTTTACTTGAAGCAAGTAAACGAAACCATGCGGTGGGCTCGAAGAGTTTCG 5226
DB 6418 AT---ACGTCATTTAATGTAATCATGCGTCAAAACCATGCGGTGGCTCGAAGAGTTTCG 6474
QY 5227 CCACCGCGCCAAAGCTGTTTGTATGACCGGCGCGGACGCTGATGCTGTATGCTG 5286

Db	6475	CGACTG	CATCGACG	CTTTT	CGAGCC	AAAAAC	CCCGT	CGACG	TGCT	CGATG	CTTT	TACGCAT	6534			
Qy	5287	GGTGT	CGTCACT	CGCAT	TGATG	TGATG	TGGG	CGAAAC	CGCT	TGGG	CGAAG	GCGG	CACG	CAGC	5346	
Db	6535	GGTGC	CGCCACT	TGCGAC	GACGT	CATTG	ACGAT	CAAA	CAC	TGTGG	GCTT	CATG	CCGAC	CAGC	6594	
Qy	5347	ATGCCG	TGCAAG	ACGCG	CAGC	ACGTAT	GACAG	CATCT	GCAAA	TCTG	CAAA	TGAAA	CCCG	CGCG	5406	
Db	6595	CCTTT	CGCAGAT	GCCTT	TGAG	CAGCG	CGCTG	TGAC	GAC	AGCTT	TGAAA	TGAAA	CGCG	T	CAGG	6654
Qy	5407	ACAGCG	GGCGG	CAATG	GATGA	ACCG	GCGT	T	TAGG	GCGT	TTC	AGGA	AGT	TGCG	ATCAT	5466
Db	6655	ACGCG	GTTG	CGCAAT	GCA	GAGC	CGCC	CTTT	TG	CCGCG	T	TTC	AGG	AGT	T	6714
Qy	5467	ACAGCT	TGCGCG	AAACAA	TGCG	GTTT	GCAT	CATCT	TGGA	AGG	CTT	CGCT	TAT	TGAT	GATG	5526
Db	6715	ATGAT	TATCG	TCCG	CTAC	CGG	TTC	GCAC	CACT	CTGGA	AGG	T	T	T	T	6774
Qy	5527	ACGAA	CATTAC	GCGAG	CTTC	GCAT	GAC	ACG	TGCG	T	TACT	CT	CT	CT	CT	5586
Db	6775	AAACG	CGCTAC	CTGAC	ACTG	ACAT	ACG	CTG	CGT	TAT	TGCT	AT	CA	CGT	CGCG	6834
Qy	5587	TCGGT	TTGATG	TGGCG	GGTAA	TGCG	GCGT	TGCG	CGAC	GAAG	CGGT	GCT	CGAT	CA	CGCT	5646
Db	6835	TGGG	CC	TGAT	GA	TGAT	GGC	CAAA	T	TAT	GGC	CGT	T	CGCG	AT	6894
Qy	5647	GCAT	TTAG	GACT	TGG	CTTC	CAG	CTCA	CT	TAA	AT	TGCG	CGCG	CACT	TGTA	5706
Db	6895	CGCAT	CTCG	GGCT	GCTT	CA	GT	TGAC	CAAC	CACT	TGCG	CGT	GA	T	T	6954
Qy	5707	AAA	ATG	CTG	CTG	CTAT	CTG	CCG	CAAT	CT	TGCG	TCAT	CA	GCGG	GGAT	5766
Db	6955	AGT	TGG	CCG	CTG	TTAT	CTG	CT	GAA	AGCT	TGG	CTGGA	AGGA	GGACT	GAC	7014
Qy	5767	CGTGA	CTGAC	CGCA	CAT	GTG	CAC	GGCT	CGCT	CACT	TGCG	ACG	CGCT	TT	TAG	5826
Db	7015	ATT	ATG	CTG	TGCG	CCAG	AAAC	CGG	CAG	CGCT	T	TAA	CGCT	AT	CGCG	7074
Qy	5827	AGG	CGG	AAAC	CCCT	AT	TAT	CACT	TGCG	CGCG	AT	CCG	GT	T	T	5886
Db	7075	AAG	CGG	AAAC	CCCT	AT	TAT	CAAT	TGCG	CGCT	CTG	GGCA	CAAT	T	T	7134
Qy	5887	CGTGG	GCAT	CGCT	TAC	GGCT	CGCG	CGCT	TTAT	TCG	CGAAA	T	TGG	CGT	CAAA	5946
Db	7135	CCTGG	CCAT	TCGCG	AC	GCA	CGA	GAC	GAGT	GTAC	CGT	TAAAT	T	TGG	CGT	7194
Qy	5947	CCGGT	GTG	CAC	CGCT	TGG	AT	TAC	GGG	CGCG	AC	CACT	TAA	AGGT	TGAAA	6006
Db	7195	CCGGT	ATA	GAG	CGCT	TGG	GAT	CA	TGCG	CAGT	CTC	AC	GTCC	ACCG	CGCG	7254
Qy	6007	TGTTGA	AGGG	CGAG	TTT	CGCG	ATCA	CTT	TGCG	TG	TGCT	CT	CTG	AAAC	CGCG	6066
Db	7255	TGCTGA	CGG	CA	TCCG	GT	CAG	CACT	T	TCTCC	GGAT	TGA	AG	CA	GTAT	7314
Qy	6067	CTGCT	CTGT	TGG	CAG	CGT	CTCT	GTG	ATTT	TAC	GTG	ACG	CTGCG	CGC	AGCG	6126
Db	7315	CTCAT	CT	TGG	CAG	CGCC	-----	GAT	CTAG	CGCAT	GCCT	TTCT	CT	CT	CAG	7368
Qy	6127	CAGCT	TAT	TAC	GGT	TCGG	GT	TAG	AGG	AAAC	CAAC	CGA	CAC	GAG	CGCT	6186
Db	7369	AAGT	TTA	GATA	AA	CGGT	CGCG	CGTA	CAGAAA	AAAC	CAAG	AGG	CA	CGC	AGCG	7428
Qy	6187	CACG	CA	TGAT	GAT	CGGT	TCGCG	CA	TGTA	TAG	CGCT	T	AG	ATAG	CGCT	6246
Db	7429	TAC	AG	CAT	GA	TG	CA	TAC	CGG	T	T	T	CAG	TAG	CGCT	7488
Qy	6247	ATAG	CGAA	CGG	CCAG	CGT	TGAT	GCAC	CAG	GGCA	CTG	TGCA	CCAT	CAAG	TAG	6306
Db	7489	GTAG	CGNA	CGG	CCAG	CGCT	GGT	GTA	CCAG	TCG	CTG	TGG	ACC	ATA	TAAC	7548
Qy	6307	GTAC	GT	CGT	CAT	TCCG	GCA	CCAA	TCCA	CTG	CAG	CGG	CCA	ATG	CGCT	6366

Db	7549	ATAAGCGGTCAATGCTTGCACAAATCCACTCGAGCGCGCAGATTCTGTGTA	CTGCCGAGTA	7606
Qy	6367	AATCAGCACCAATCCGCCAGTAGTCCGCAAAACACACCCGCAATGAAGATCGTTGAGCTCAAACTT	6426	
Db	7609	AATCAGGGCAATCGACACATGGCGAATACCAACCGCATAGAGATCGTTAACTTCAATGC	7668	
Qy	6427	ACCCTGTGCGGTTTCATGGTGCAGACAGATGCCAGAGCCCATCCCCAACCGGTGCATGATGTA	6486	
Db	7669	GCCTTTACGCGGTTTCATGATGTGAAGATGCCAGAGCCCAACCCCGTGCATGATGTA	7728	
Qy	6487	TTTATGCGACAGCGCCCTACGATTTCATCCACACCGTTGCGCAACAAGATAGCAC	6546	
Db	7729	TTTATGTCGCGAGTGCAGCAACCACTTCCATGCGCGACCAACCGTGCACAAACGATCAGGC	7788	
Qy	6547	GTTCCTAAACACAGACGATGTTGCTCCA	6574	
Db	7789	ATTCCAATCCACACATATTTCTCAA	7816	
RESULT 4				
ID	ADQ77262	standard; DNA; 8609 BP.		
AC	ADQ77262;			
DT	23-SEP-2004	(first entry)		
DE	Nucleotide sequence of plasmid pPCB15.			
KW	carotenoid, carotenoid enzymatic biosynthetic pathway; dxs gene;			
KW	D-1-deoxyxylulose 5-phosphate gene; idi gene; ygbBP gene; ispB gene;			
KW	yJgR gene; oligo-ribonuclease gene; dxs gene; dxr gene;			
KW	DOXP reductoisomerase; yggP gene; ispD gene; ycbB gene; ispE gene;			
KW	ygbA gene; ispF gene; lycB gene; ispH gene; idi gene;			
KW	isopentenyl diphosphate isomerase; ispA gene;			
KW	farnesyl pyrophosphate synthase; ispB gene;			
KW	octaprenyl diphosphate synthase; crtE gene;			
KW	geranylgeranyl pyrophosphate synthase; crtB gene; phytoene synthase;			
KW	crtI gene; phytoene dehydrogenase; crtX gene;			
KW	zeaxanthin glucosyl transferase; crtZ gene; beta-carotene hydroxylase;			
KW	crtY gene; lycopene cyclase; enzyme; ss.			
OS	Synthetic.			
XX				
PN	WO2004056975-A2.			
XX				
PD	08-JUL-2004.			
XX				
PF	19-DEC-2003; 2003WO-US041812.			
XX				
PR	19-DEC-2002; 2002US-0434618P.			
XX				
PA	(DUPO) DU PONT DE NEMOURS & CO E I.			
XX				
PI	Cheng Q, Rouviere PE, Suh W;			
XX				
DR	WPI; 2004-525437/50.			
XX				
PT	Novel carotenoid overproducing bacteria such as Escherichia coli or			
PT	Flavobacterium, comprising genes encoding functional carotenoid enzymatic			
PT	biosynthetic pathway, useful for producing carotenoids such as			
PT	astaxanthin and beta-carotene.			
XX				
PS	Disclosure; SEQ ID NO 64; 137pp; English.			
XX				
CC	The specification describes a carotenoid overproducing bacteria, which			
CC	comprise genes encoding a functional carotenoid enzymatic biosynthetic			
CC	pathway, where the dxs (D-1-deoxyxylulose 5-phosphate gene), idi, ygbBP			
CC	and ispB genes are overexpressed, and the yJgR gene (oligo-ribonuclease			
CC	gene) is down regulated. In bacteria of the invention, the carotenoid			
CC	enzymatic biosynthetic pathway consists of the genes dxs, dxr (DOXP			
CC	reductoisomerase), yggP (also known as ispD), ycbB (also known as ispE),			
CC	ygbB (also known as ispF), lycB (also known as ispH), idi (isopentenyl			

CC	diphosphate isomerase), ispA (farnesyl pyrophosphate synthase), ispB (octaprenyl diphosphate synthase), crtB (geranylgeranyl pyrophosphate synthase), crtB (phytoene synthase), crtI (phytoene dehydrogenase), and crtX (zeaxanthin glucosyl transferase), crtZ (beta-carotene hydroxylase) and crtY (lycopene cyclase). The lypB and dxr gene are optionally overexpressed. Carotenoid overproducing bacteria of the invention are useful for producing carotenoids. The present sequence represents a CC plasmid pPCB15, which is used during construction of bacteria of the CC invention.
XX	
SQ	Sequence 8609 BP; 1918 A; 2325 C; 2309 G; 2057 T; 0 U; 0 Other;
	Query Match 37.4%; Score 2619.6; DB 12; Length 8609;
	Best Local Similarity 64.7%; Pred. No. 0;
	Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;
QY	373 CAGCTGACGCAAAATTTCAGCGGATCTTTGAACATTTTACTGCTGCGGACAGCAAAAGC 432
DB	1627 CAGTTGCTGGTGATATCGATAGCGCTTGATCAGTTACTGCGGTTTCAGGTTGAGCGG 1686
QY	433 GATCGCGTGGTCCGCGATGCGTCCGGAACGCTGGCGGACAGGCAAGATTTCGTCT 492
DB	1687 GATTGTGTGGTGGCGGATGCTGAAGGACGCTGGCACCGGCAAAACGATTTCGTCCG 1746
QY	493 TTATTACTGCTGTGCGACGCGCGATATGGTTGCGAGCTCACGCAAAATGGCGTTCTC 552
DB	1747 ATGCTGCTGTTATTAAACGCGCGGATCTTGCTGTGCGATCAGTCACGGGGATTA 1806
QY	553 GATCTGCGTGTGCTAGTGAATATGTTGACGCGGATCGCTGATTCTTGATGACATGCC 612
DB	1807 GATTTAGCTGCGCGTGTGAATATGTTGATGCTGCTGCTGATTCTGATGATATGCC 1866
QY	613 TCGATGATTAACGCGCAGATGCTGTGTGCTGCGCTACCGTGCATCGCAATTTGGTGA 672
DB	1867 TGCATGGAGATGCGCAGATGCGTGGGGGCGTCCACCATTCACACGCGAGTACGGTGA 1926
QY	673 AAGCTGCGGATCTCGCGGCAATCGCGCTGCTTAGCGCGGCAATTTGAAGTGAATGCCAT 732
DB	1927 CATGTGCGGATCTGCGGGGCTGCTTTACTCAGCAAAAGCGTTTGGGTGATTCGGAG 1986
QY	733 GCACCGGTTGCTGCTGCTATATAAATCTGAAGCATGCTGTAATCTCCGCTGCGCTC 792
DB	1987 GCTGAAGGTCTGACGCGGATAGCAAACTCGCGCGGTGTGCGAGCTGTCCAATCGATT 2046
QY	793 GGCCTGCGGGCTTAGTGAAGGCAATTCAGATCTGCAAGCGGACGCGAGAGCCCGC 852
DB	2047 GGCATGAGGCTGTTGTTAGGCGCAGTTTAAAGACCTCTCGAAGGCGATAAACCCCG 2106
QY	853 AGCCCGAAGCGATCGCATGACCAACGAACTGAAACACAGCGTGTGTTTCGCGCACG 912
DB	2107 AGCGCGGATGCCATCTGCTTAAACCAATCAGTTTAAACACAGACGCTGTTTTCGCGTCA 2166
QY	913 CTGCAATGCGGGGATTCGCTGAGCTTACCGCAGTGGCGGCAAAAGCTTAGCTTC 972
DB	2167 ACGCAATGCGCTCAATTGCGGCAACGCGTCTCGAAGCGGTGAGAACCTTGCAATCGT 2226
QY	973 TTGCGCCAGGATTTGGCGCAGGCGTTTCAACTGCTCGACGACCTCGCGGACGTTGCAAA 1032
DB	2227 TTCTGCTGATCTCGCGCAGGCTTTCAGTTGCTTGAACGATCTTACCGATGGCATGACC 2286
QY	1033 CACACCGGTAAGATGTGACAGGATCAGGGCAAAATCAGCGTGGTACAGATGCTCGGT 1092
DB	2287 GATACCGGCAAGACATCAATCAGGATGCAAGTAAATCAACGCTGGTCAATTTATTAGGC 2346
QY	1093 GCTGACCGCGGAAAGCTGCGCTGCGGATCACTGCGGACGCGAGATGACACCTTGCC 1152
DB	2347 TCAGCGCGTTCGAAGAACGCTTCGACAGATTTGCGCTGCGCAGTGAACCTTTCC 2406
QY	1153 TGCCTCTGCCATCGCGCATGCGCACTCGCCAA---TATTAGCAGCGCTGTTTATCAA 1209
DB	2407 GCGGCATGCCAAACCGGCAATTCACACCAACCACTTTTATTAGGCGCTGGTTTGACAA 2466
QY	1210 CAGCTAGGATATTCAACTGAGCGGCTCAGCCGCTGGGCCACTTTTTCGCTGATCGCGC 1269

DB	2467 AACTCGCTCGCTCAGTTAAGGATGCTGCA-----TGAGCCATTTTGGGTGATCGCAC 2521
QY	1270 CGCGCTCTACAGCACTTTACAGCGTTGACAGCGTTAGCAAAACGCTGCTGGCGCGG 1329
DB	2522 CGCCCTTTTTCAGCCATGTTTCGGGCTCTGCAAAACCTTGTCTAGGAATTAGTGGCGCGG 2581
QY	1330 GCATCGCATCACATTCATCCAGCAAGCGATCCCGCACTTTTGTCTTAGCGAGAAACGCA 1389
DB	2582 GTCATCGTTGTACGTTTTTTCAGCAACATGACTGCAAGCGCTGGTAAACGGGACGATA 2641
QY	1390 TCGATTTTGTGCGCTCGGCCAAACAGACGATCCTGCGGTTTCGTGCGCGCCGCTGTGTC 1449
DB	2642 TCGATTTCCAGACCGCTCGGACTGCAACGCACTCTCCCGGTTCTTATTCGCACTGCTGC 2701
QY	1450 ATCGGCTGGCTCGCGGCGGCTGCTGCTGTTTTCGCGTGATCGAGATCTCGCGCTCT 1509
DB	2702 ACCTGGCGCGCAACCCACTCGGACCTCTGATGTTTACGACTGATCAATGAATGCGACGTA 2761
QY	1510 GCACGATATGCTGTGCGCGAACTGCTCGGCTACTGAAAGCATTTGAAACATCGATGGCG 1569
DB	2762 CCAGCATATGCTTTGCGGGAACTGCGCGGCTTTTCATGCTTGCAGATAGAGGCG 2821
QY	1570 TGATCGCGCAGAAATGGAAGCGGCGGCGGATTTGCTGCTGAAGCGCTGCATCTGCGCT 1629
DB	2822 TGATCGTTGATCAAAATGGAGCCGCGAGTGCAGTAGTCGAGAGCGCTCAGGCTCGCGCT 2881
QY	1630 TTGTTTTCGCTGCGCTGCGCTTTCGCGTCAATCGTGAAGCGGATTTCCGTTTTCGCGTGA 1689
DB	2882 TTGTTTTCGCTGCGCTGCGCTGCGCTGCGCTCAACCGCAACCGGTTTTCGCTTTCGCGTGA 2941
QY	1690 TGCCCTTTCGCTTTTTCACAGGATGACAAAGCGCTGAAAGCTTTTTCAGGCGCAGCAGGATA 1749
DB	2942 TGCTTTTCGAGTACGCGACACAGCGATGCGGCTCGGAAACGCTATACCAACCGCGGAAAAA 3001
QY	1750 TCTATGATCGCATCATGCTGCTCACCGGACGCTGATCCTCAACACGCGCGGCTGTTA 1809
DB	3002 TTTATGATGCTGATGCGAGCTCAGCATCGTGTGATCGCGCATCATGATGCGAGATGG 3061
QY	1810 ATTTGACGCGAGCGCGGATTTACATCAGTGCCTGTGCGCGCTCGCAAAATACGACGAGA 1869
DB	3062 GTTTAGCCCCGGTGAATAAATGTCATCATTTTCTTCTCCACTGCGCAAAATCAGCCAGT 3121
QY	1870 TGGTGGCGGCTTTGATTTTTCAGTCAAGTCAACGCGCGCTGCTATCACCGCTGCGGCG 1929
DB	3122 TGATCCCGAACTGGAATTTTCCCGCAAGCGCTGCCAGACTGCTTTTCATCGGTTGAGC 3181
QY	1930 CACTCGCG---CCCGGTTTCTCTGCGCGCTCCATGCGCCCTTGGCCAGCGCTCGCTC 1986
DB	3182 GCTTACGCAACCCAGGGAGCGCGGCTCATCAACTTCTTATTTTCCGTCGCCGACA 3241
QY	1987 AGCCCGTGGTTTATGCTCGCTGGGTACGCTGCAAGGCGCATCGCTTCGCGCTGTTCTGC 2046
DB	3242 AACCCGCTATTTTTCCTCGCTGGGCACTTCGAGGACATCGTTATGCGCTGTTTTCAGGA 3301
QY	2047 ATCTGCGCAGCGCTGCGCGCAGCTGCGCTATCGCTGATCGCCCATTTGTGGGGAT 2106
DB	3302 CCATCGCAAGCGCTGCAAGAGGTGGATGCGAGTTACTGTTGCGCACATCTGTGCGGCGC 3361
QY	2107 TAAACGCGGAAACAGACGATCAGCTGGAGCTGCTGGCGCGCGCGTGGGTGACGATTCG 2166
DB	3362 TCTCAGCACGCGCAGCTGAATGCTGGCCCGGGCGGGAATTCAGGTTTGTGGATTTTG 3421
QY	2167 TCGATCAGGCGCGCGCTTACAGCAGCGGCTGTTTATCATCTCATCGCGGTTTAAACA 2226
DB	3422 CCGATCAATTCGCGACACTTTCACAGGCACTTCAAGTCAATCACATGTTGGGATGAATA 3481
QY	2227 GCGCTGGAAGCACTGGAATGCGGTAGCGGATGCTGGCGCTGCGGATGCTGTTTGTGATC 2286
DB	3482 CGGTACTGACGCTATTTGTTTCCCGCACACCGCTACTGCGCTGCGCTGGCATTTGATC 3541
QY	2287 AGCCCGGCTGCGCGCGCATTTAGTGGCATGAGTTGGTTCGCGCGCATCACGCTTTA 2346

Db 3542 AACCTGGCGTGCATCACGAATTGTTTATCATGGCATCGCAAGCGTGCCTCTCGGTTTA 3601
Qy 2347 GCGGTGTTTATCAACTGGAGACGATCTGCAACAGCTGTGACCGACGATCGTTACGCGC 2406
Db 3602 CTACAGCCATGCGCTGGCGCGCAGATTGATCGCTGTGACTAAACCGATTACCCGC 3661
Qy 2407 TACGGATGTCAGGATTCAGGCGCAGCTGACGGCGCAGCGGTTGCCAGCGTGCAGCG 2466
Db 3662 AGCGTATGACAAAATTCAGGCGCGCATGCGTCTGGCAGCGCGCACACAGCGCGCGCG 3721
Qy 2467 ACATCGTGCAGAGCGGCTGTGACAGCAGCAAGTCTGCTGGCGGAGCGACCTGATGG 2526
Db 3722 ATATTGTTGACAGGCGGATGGG-GACTGTGACGCACTACTCAGTGGCGAGATTATGCA 3780
Qy 2527 CACGCAATACGATGTTGTTGCTGGTCTGGAATGCGGCAATGCGTTCGATGCGCTGCG 2586
Db 3781 ACCGCACTATGATCTCACTTCTGCTGGTGGTGGTCTGCTAATGGCTTATGCGCTCG 3840
Qy 2587 TCTGCGTCAATTGCAACCACTGAATGAATGCCCTGTTGCTGGAGAGCGATCGCATCCGC 2646
Db 3841 GCTTCAGCAACAGCATCCGGATATGCGGATCTTGGCTTATGAGGCGGCTCCTGAGCGGG 3900
Qy 2647 AGGCAATCATCTGCTGCTGTTTATCACAGCGATCTCAGCGCGAACAACCTCGCTGGCT 2706
Db 3901 AGGGAACCATACCTGCTCTTTCAGAGAGGATTTAACTGATCAGCATCGCTGGAT 3960
Qy 2707 GCAACCGCTGATTTACCGTGGTGGTTCAGGTTATCAGGTGCGTTTCTGCGCTGCGCG 2766
Db 3961 AGCGCGCTTGTGCTCATCTGCGCGGACTACCAAGTTGCTTCCCGCAACGCGCTCG 4020
Qy 2767 CAATCTGGAGCGGATTTATGCTCATCGCATCAGCGGATTTTCCCGCGCATCTTTACGC 2826
Db 4021 CCATGTGAACAGTGGCTACTACTGCTGACCTCCCGGCATTTGCGCGGATATCTCGGCA 4080
Qy 2827 GCGATGGGTGACGATCTGTGCAACACAGCGCTACACAGGTAAACCCACGCGAGT 2886
Db 4081 ACAGTTTGAACACATTTATGCTGATACCGCGTTTTCAGCGTTTCATGCTGAATCGGT 4140
Qy 2887 GACGCTGGCGGATGGCGGTGAATCTGCTGCGAAGTGGTGAATGATGGTTCGCGCGCTGCA 2946
Db 4141 CCAGTTAGCGGATGGCGGATTTATCCAGTACAGTATGATCGAGCGGAGGGTTTACAC 4200
Qy 2947 GCGGAGCCACATCTCGAGTGGTTTATCAGGTGTTTCTTGACAGAGTGGCAGTGGC 3006
Db 4201 GCTGATTTGCACTACGCTAGGATTCAGGCAATTTATCGGTGAGGAGTGGCACTGAG 4260
Qy 3007 GCAGCGCGCGCTGACAGCGGATCTGATGATGCCACCGTGCATCAGCAAGCGGG 3066
Db 4261 CGGCGCGCATGTTTATGCTACCGATATCATGGATGCAACGCTGCATCAGCAAAATGG 4320
Qy 3067 TTATCGTTTTGCTACACGCTGCGCTCAGCGCGGATCGGCTATTGATTGAAGATACCCA 3126
Db 4321 CTACCGCTTTGTTTATACCTGCGCTTTCGCAACCGCACTGCTCATCGAAGACACACA 4380
Qy 3127 TTACGTTAACAGCGCGCTGGCGGAGAACACCGCTCTGTGACGACATCGCGGACTATGC 3186
Db 4381 CTACATTGACAGGCTTAATCTTCAGGCGGAAACGCGCGCTGCAAACTTCGCGGATTATGC 4440
Qy 3187 CAATCAGAGGCTGACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3246
Db 4441 TGGCGACAGGTTGGCGGTTACAGACGTTGCTGCGGAGAACAGGGTGCTATGGCCAT 4500
Qy 3247 TACCTCTGAGCGGCAACATCGATCGATTCTGGCAACAGCAGCGCGGCAAGCGTGCAGCG 3306
Db 4501 TAGTTAACCGGCGAATACTCGTCAGTTTGGCAACAGCAAC--CGCAAGCTGTAGCGG 4557
Qy 3307 CTTGCGCGCGGCTGTTTATGCGCACCGGTTTACTCTTCTGCGCTGCGCGCTGCGCT 3366
Db 4558 ATTACGCGCGGCTGTTTATCCGACACCGGCTACTCTCTTACCGCTGCGCGTGGCGCT 4617
Qy 3367 AGCGGATGTTGAGCAGCGCTGTTGCCACCGATGCCCTCAGCTCAGCGCAACATATCGA 3426
Db 4618 GGCCGATCGTCTCAGCGCGCTGATGTTTACCTCTTCTCTGTTTACACAGCAGATTGC 4677

Qy 3427 AGCTTTGCGCGCTCAGCAGTGGCGGAACAGCGAATTTTTCGCTGCTAAACCGCATGCT 3486
Db 4678 TCACITTTGCCAGCAACGTTTGGCAGCAACAGGGGTTTTTCCGCTGCTGAATCGCATGTT 4737
Qy 3487 GTTTTGGCGCGTAAAGCCGACAGCGCTGGCGGTGATGCAACGTTTTTACCGGCTCGA 3546
Db 4738 GTTTTGGCGGACCGCGGAGTCAAGCTGGCGTGTGATGCAAGCGTTTCTATGGCTTACC 4797
Qy 3547 TGCGGGGTTAAATTAGCCGCTTTTACGCGGGCAACTGCGCCCTGCGGATATAAACCGCGAT 3606
Db 4798 CGAGGATTTGATTTGCCGCTTTTATGCGGGAACCTCACGCTGACCGATCGGCTACGAT 4857
Qy 3607 TCTGTGGGCAAGCGCGCTGCCCATCGGTGAAGCGCTGCGCGCTCTGTGAATCTGT 3666
Db 4858 TCTGAGCGCAAGCGCGCTTCCGCTTTTCGCGCATTTGCAAGCAATTTATGACGACTCA 4917
Qy 3667 CGAACCGGGAAGAAAATGAACGCACTTATGATTTGGCGCAGGCTTTGCGCGCTG 3726
Db 4918 TCGTTGAAGAGCGACTACATGAACCACTACGTTAATTTGGTGGCGCTTTTGGTGGCTG 4977
Qy 3727 GCGCTGGCGATTCGCTGCAAGCGCGGCGCATACCAACCACTTACTCGAGCAGCGCGAC 3786
Db 4978 GCACTGGCAATTCGTTTACAGCGCGCAGGATTTCTGTTTTGCTGCTTTCGAGCAGCGGAC 5037
Qy 3787 AAACCGGCGGACGCGCTTATGTTTGAAGACAGTGGCTTTTACCTTCGATGCGGACCC 3846
Db 5038 AAGCGGGTGGCGGCTTATGTTTATCAGGAGCAGGCTTTTACTTTTATGATGCGGCTT 5097
Qy 3847 ACCTGATCACCAGTCCAGCGCCATCGAAGAGTTGTTTACGCTGCGAGGAATCGCTC 3906
Db 5098 ACCGTTATCACGATCCAGCGGATTTGAAGAACTGTTGCTTGGCGGTTAAACAGCTT 5157
Qy 3907 ACCGATTTAGCTGAGCTGATGCGGTAAACGCTTCTATCGCTGCTGCTGGGAAGATGGC 3966
Db 5158 AAGGATTTAGCTGAGCTGTTGCGGTGACGCGGTTTTATCGCTGCTGCTGGAGTCCGGC 5217
Qy 3967 AAACAGCTTGATTACGCAATTAATCAGCGCTGCTGGAGCAGAGATCGCCAGTTCAAT 4026
Db 5218 AAGGCTTTCAATTTAGCATAAACGACAGGCCAGTTTGAAGCGCAGATACAGCAGTTTAAAT 5277
Qy 4027 CCGCAAGATGTAGAAGCTATCGTCAATTTCTTGGCTATTCACGTTGAATTTTACAGAG 4086
Db 5278 CCGCGGATTTGCGGGTTATCAGCGTTCTTGAATTTTCGCTGCTGCTGCTTCAATGAG 5337
Qy 4087 GGTATCTGAAACTCGGACGCTGCTTCTGCAAGTGGCTGCAATGCTGCGCTGCGC 4146
Db 5338 GGTATCTGAACTCGGACGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5397
Qy 4147 CCGCAGTTGGGAGCTGTGCAAGCATGGCGCAGCGTCTACAGCATGGTGGCGAAATTTAT 4206
Db 5398 CCCCAGTTGGCAAGCTGCAGGCATGGCGCAGCGTTTACAGTAAAGTTTTCGCGCTACAT 5457
Qy 4207 GAGGAGCATCTGGCTGAGCGGTTTTCTTCCACTCATCTGCTGCTGCGGCTATCCT 4266
Db 5458 GAGGATGAGCATTTCTGCGCAGCGTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 5517
Qy 4267 TTTGCAACGCTCATCTATACCTTAAATTCATGCGCTGAGCGTGAATGGGCGCTGCTG 4326
Db 5518 TTTGCAACCTCTCTTATACGCTGATTACGCGTTAGAACGGGAATGGGCGCTGCTG 5577
Qy 4327 TTTTCCGCGCGGCAACCGCGCTGTTGAGAGGCGATGGCGGCTGTTTCGAGGACTTGG 4386
Db 5578 TTTTCCAGCGGTTGAACCGCTGCTGCTCAATGGCATGATCAAGCTGTTTTCAGGATCTG 5637
Qy 4387 GCGCGGAGCTGTTTACTGATGCGGAGTGAAGCTGCGAAGCTGGAACCGCGCATCCGAT 4446
Db 5638 GCGCGGAGCTGCTGCTTAAACCGCGGCTCAGTCAATGGAACCGTTGGGCAAGATT 5697
Qy 4447 AGCGGCTTCACTGATAGAGGCGGACGACGCTTTCGATGCGCGCTGCTGCTTCAATGCC 4506
Db 5698 CAGGCGGTCAGTTGGAAGACGCGCAGACGCTTTTGAACCTGCGCGTGGCGTTCGAGCT 5757

QY	4507	GAGCTGGTGATACCTACGACAAACTGCTTCCGCCACCATCCGCTGGCAATGAAA	CGTGGC	4566
Db	5758	GATGTTGTATACATACCTTATCGCGATCTGCTGTCTCAGCATCCCGCAGC	CGCTAAGCAGCG	5817
QY	4567	ACATCCGCTGAAGCGTAAAGCGCATAGACAACTCGCTGTTTGTACTCTATTTTGGCCCTGGAAT	4626	
Db	5818	AAAAAATCTGCAATCCAAGGATAGATTAATCACTGTTTGTACTCTATTTTGGTCTCAAC	5877	
QY	4627	CAGCCGCATGAACAGCTCGCGCACCAACACCGTCTGTTTGGCCCGCGTTTATCGTGA	GTG 4686	
Db	5878	CATCATCAGATCAAACTCGCCCATCATACCGTCTGTTTGGGCGCACGCTACCGTGAAC	CTG 5937	
QY	4687	ATCGATGAGATTTTCAACAGCAGCCAGCTGGCAGACGATTTTTCATCTTACCTTGCACCGG	4746	
Db	5938	ATTCAAGAAATTTTAAACCATGATGTCTGGCTGAGGATTTTTCGCTTTATTTTACAGCA	5997	
QY	4747	CCCTGCAGCAGCATCCGCTGCTGGCACGCCCGCTCGCGCAGCTTTTATGTGTTAGCG	4806	
Db	5998	CCTTGTGTACGGATCCGTCACCTGCGCACCCGAAGGCTGCGCAGCTATTATGTGCTGGCG	6057	
QY	4807	CGGTCGCGGATCTCGGCACCGCTGACATCGACTGCGAACAGGAAGGACCGGCTTGGCG	4866	
Db	6058	CCTGTTTCCACACTTAGCGACCGGGAACCTCGACTGGCGCGTGAAGGACCCGACTGCGC	6117	
QY	4867	GATCGAATTTTTCCTTATCTGGAAGACGACTACATCCCGGATTAGCTCAGCAATTAGTG	4926	
Db	6118	GATCGTATTTTGACTACCTTAGCAACATTACATGCTGGCTTGGGAAGCCAGTTGGTG	6177	
QY	4927	ACACACAGAAATGTTTACGCGTTTGAATTTTCGGGACACGCTGCAATCGCCATACGCGCTG	4986	
Db	6178	ACGCACCGTATGTTTACGCGGTTTCGATTTTCGGCGACGAGCTCAATGCCCTGGCAAGGTT	CG 6237	
QY	4987	GCCTTTTCGCTGAGCGCGATTTTGAAGCAAGCGCTGTTTCCGCCCGCATACCCCGAT	5046	
Db	6238	GCCTTCTGGTTGAACCTATTCTGACCCGAGCGCTTGTTTCGACCAATACCCGAT	5097	
QY	5047	GCCGATATCAGCAATCTTATCTGTTGGTGC	CGGTA CGCATCCAGCGCGGCGTGCC 5106	
Db	6298	AAGCACATTGATAATCTTTATCTGGTTGGCGCAGGACCCATCTCTGGCGCGGCATTC	CG 6357	
QY	5107	GGCGTGATCGGTTTGGCGAAGGCCACCGCCAGGCTGATGCTGAGAGATCGCGCCGAATGA	5166	
Db	6358	GGCGTAATCGCTCTGGCGAAGGCGACCGCAGGCTTAATGCTGGAGAGCTTGATTTGACGA	6417	
QY	5167	ATCGACAGCCTTTACTTTGAGCAAGTAACGCAAAACCATGGCGTGGGCTCGAAGATTT	CG 5226	
Db	6418	AT - - - ACGTCAATTACTGAATCATGCCGTGAAACCATGGCGTTGGCTCGAAAAGCT	TTG 6474	
QY	5227	CCACCGCGCAAGCTGTTTGATGCACCGAGCGCGCGACGACGCTGATGCTGTATGCGT	5286	
Db	6475	CGACTGCATCGACGCTTTTCGACGCCAAAACCCGCTCGACGCTGCTGATGCTTTACG	CGAT 6534	
QY	5287	GGTGTCTGCTACTGCGATGATGATGATGGGCAAAACGCTCTGGCGAAGCGCGCACCG	CAGC 5346	
Db	6535	GGTGCCGCACTGCGACGACGTCATTGACGATCAACACCTGGGCTTTCATGCGCAGC	AGC 6594	
QY	5347	ATGCCGTGGAAGACGCGACGACGATGAGCATCTGCAAAATGAAACCCGCGCGCCT	5406	
Db	6595	CCTCTTCGAGATGCCCTGAGCAGCGCCTCGACGAGCTTGAATGAAACCGCTCAG	CGCT 6654	
QY	5407	ACAGCGCGCGCACATCGATGAACCGGCGTTTATGGGCGTTTTCAGGAAGTGGCGAT	CATT 5466	
Db	6655	ACGCCGTTTGCAAATGCAAGACCGCTTTTTCGCCGCTTTCAGGAGTTCGCAATGG	CGC 6714	
QY	5467	ACCAAGTCCCGCAACAACTGGCTTTGATCATCTGGAAGCGTTTCGCTATGATGATG	CAACGCA 5526	
Db	6715	ATGATATCGTCCCGCCTACGCGTTCCACCATCTGGAAGGTTTTCGCCATCGATG	TCGCG 6774	
QY	5527	ACGAACTTAGCGAGCTTCGATGACACGCTGCGTTACTGCTATCACTGCTCGCGG	CGTGG 5586	
Db	6775	AAACGCGCTACCTGACACTGACGATACGCTGCGCTTATGCTATCACTGCTCG	CGGTTG 6834	
QY	5587	TCGCTTTGATGATGGCGCGGCTAATGGCGCTGCGCAGGAAGCGGTGCTCGATCA	CGCCT 5646	

Db	6835	TGGCCCTGATGATGGCGCAAAATTATGGCGGTTTCGGATACGCCACGCTCGATCGCGCT	6894
Qy	5647	GCATTTTAGGACTGGCGTTCCAGCTCACTAACATTCGGCGCGACATGTAGAAAGATGCCG	5706
Db	6895	GCATCTCGGCTGGCTTTCCAGTTGACCAACATTCGCGGTGATATTGTTCGACGATGCTC	6954
Qy	5707	AAAATGTCCTGCTATCTGCGCAATCTCGGCTCGATCAGCGGGATACGCCCGGATA	5766
Db	6955	AGTGGCCCGCTGTATTCTCGCTGAAGCTGGCTGGAAAGGAAGGACTGACGAAGCGA	7014
Qy	5767	CGCTGACTGCACCGCAACATCTGTGCAGCGCTCGCCCTCACTGGCAGCGCGTTTGTAGTCGG	5826
Db	7015	ATTATGCTCGCCAGAAAAACGGCAGCGCTTAAGCCGTATCGCGCGGAGCTGTACGGG	7074
Qy	5827	AGCGGAAACCGTTATTATCACTCGCGCGGATCCGGTTTACGGGTTTTACCGCTCGCTCGG	5886
Db	7075	AAGCGGAACCGCTATTAGTGATATCATATGCGCCGCTCTGGCACAATTAACCTTACGTCGG	7134
Qy	5887	CGTGGGCCATCGCTACCGGCTCGCGCGGTTTATCGCAAAATGGCGCTCAAAAGTTTTCAGCAG	5946
Db	7135	CTGGGCCATTCGCGACGACGACGAGCGTGTACCGTAAATTTGGGTGAAAGTTGAACAGG	7194
Qy	5947	CCGCTGTGCACGCTCGGATTTACGCGCAGCGCACACGTAAGGTGAAAACTCGCGCTGC	6006
Db	7195	CCGCTAAGCAGGCTCGGATCATCGCCAGTCCACGCTCCACCGCGCAAAAATAACCGTTT	7254
Qy	6007	TGCTGAAGGGCAGGTTTCGGCATCACTTCGCGGTGTCTCGTCTCTGAACCGCGTCGCG	6066
Db	7255	TGCTGACGGGCATCCGGTCAGCGAGTTACTTCCCGGATGAAGACGTATCCACCCGCTCTG	7314
Qy	6067	CTGCTCTGTGCGACGCTCTCGTTGATTTTACGTCCTGTGACGCTGGCGCAGCGTGGCTG	6126
Db	7315	CTCATCTCTGCGACGGCCC-----GATCTAGCGCGATGCTTCTCTCAGCGTGCCTG	7368
Qy	6127	CAGCTTATTACGGGTGGCGGTAGAGAAACCAACACGACGCGCTTTCACGCCCGCG	6186
Db	7369	AAGTTTAGATAACCGTGGCGGTACAGAAAACCAAAAGGACACGCGAGCCCTCTTTTCCCT	7428
Qy	6187	CACCGCATGATCGGCTGGCGCATATGAAGCGCTTAAGATAGAGCTTTGGCGCGGAT	6246
Db	7429	TACAGCATGATGATACGCTGGGCCATGTATACCGTTTCAGGTAGCTTTTGGCGGTAT	7488
Qy	6247	ATAGCGAAACGGCAGCGTTGATGCAACAGCGCATCTGTGACCATGAAGTAGAGCGGCC	6306
Db	7489	GTAGCGAAACGGCAGCGCTGGTACAGTCCGTCGTGGACCAATAAAATACAGTAAACC	7548
Qy	6307	GTAGTCTGTCAATCCGCGACCAATTCACATGCGAGGGCCACATGCTTTCACACCGCAT	6366
Db	7549	ATAAGCGGTATGCTGCAACCATTCATCTGGAGCGGCAGATTCTCTGTACTGCGGAGTA	7608
Qy	6367	AATCAGCAAAATCGCCAGTACCGCAAAACACACCGCATAAAGATCGTTGTAGCTCAAACTT	6426
Db	7609	AATCAGGSCAATCGACACAATGGCGAATACCAACCGCATAGAGATCGTTTAACTTCAATGC	7668
Qy	6427	ACCGTGTGCGGTTTCATGGTTCGACAGATGCGACGCCCTCCCAACCGTGCATGAGTA	6486
Db	7669	GCCTTTTACCGGTTTCATGATGTGAAGATGCGACGCCCAACCCAGCGCGTGCATGATGTA	7728
Qy	6487	TTTATGCGACAGCGCGCTACGATTTCCATCACCAACCGGTTGCCAACAGATGAAGCAC	6546
Db	7729	TTTATGCGAGTGCAGCAACCACTTCCATGCGCACCGACCGGTGACAAACACGATCAGGGC	7788
Qy	6547	GTTCCATAACCGAGCATTTGTTGTTCCA	6574
Db	7789	ATTCCAAATCCACAACATAATTTCTCAA	7816

RESULT 5
ADQ48673
ID ADQ48673 standard; DNA; 8609 BP.
XX AC ADQ48673;

QY 1810 ATTTCAGGAGCGGCGGATTACATCAGTGCCTCTGCGCGTGCACAAATCAGCCAGA 1869
DB |||||
QY 3062 GTTTAGCCCCGGGAAACATGCAATCTGTTTCTCCACTGGCAAAATCAGCCAGT 3121
DB |||||
QY 1870 TGGTGGCGGCTTTGATTTTCACGTCAGCAATCGCCGCTGCTATCACGCGCTGGGGC 1929
DB |||||
QY 3122 TGATCCCGAATGATTTTCCCGAAGCGCTGCCAGCTGCTTTTATGCGGTGGAC 3181
DB |||||
QY 1930 CACTCCGCG- - -CCCGGTTTCTCTGCGCGCTCCATGCGCCCTGGCCAGCGCTGCCTC 1986
DB |||||
QY 3182 CGTTACGGCAACCCAGGGGAGCGCGGGTCAATCAATCTTTATTTTCGTCGCCGACA 3241
DB |||||
QY 1987 AGCGGTGTTTATGCTGCTGGTTCAGTTCGTCGAGGCAATCGCTTCGCGCTGTTCTGC 2046
DB |||||
QY 3242 AACCCTGATTTTTCGCTCGCTGGGACCCCTCAGGGACATGTTATGCGCTGTTTCAGGA 3301
DB |||||
QY 2047 ATCTGCGCAGCGTGCGCCAGCTGCGGCTATCGCTGATGATCGCCCATTTGTGGGGAT 2106
DB |||||
QY 3302 CCATCGCCAAAGCTGCGAAGGTGGATGCGCAGTTACTGTTGGCACACTGTGGCGGC 3361
DB |||||
QY 2107 TAAAGCCGAAACAGACGCAATCAGCTGAGAGCTGCTGGCGCGCTGGGTGACGGATTTGC 2166
DB |||||
QY 3362 TCTCAGCCACGAGGAGGTGAATGCGCCGCGGGGAGACATTCAGGTTGTGGAATTTG 3421
DB |||||
QY 2167 TCGATCAGCGCGCAGCCCTACAGCACGCGCAGCTGTTTATCACTCATGCCGGTTAAACA 2226
DB |||||
QY 3422 CGGATCAATCCGAGCAGCTTTTACAGGCACAGTTGACATCAACATGTTGGATGAATA 3481
DB |||||
QY 2227 GCGCGCTGGAAGCAGCTGGAATGGGTAACCGATGCTGGCGTCCGATTTGCTTTGATC 2286
DB |||||
QY 3482 CGGTACTGAGCGCTATTGCTTCCGCAACACCGCTACTGCGGCTGCGCTGGCATTTGATC 3541
DB |||||
QY 2287 AGCCCGGCTGGCGCGCGCATTTAGTGGCATGACGTTGGTGGCGCGCATCACGCTTTA 2346
DB |||||
QY 3542 AACCTGGCTGGCATCACGAATGTTTATCATGGCATCGGCAAGCGTGGCTTCGCTTTA 3601
DB |||||
QY 2347 GCGGTGTCATCACTGAGGAGCAGTCTGCAACAGCTGCTGACCGACGATCGTTACGCGC 2406
DB |||||
QY 3602 CTACAGCATGCGCTGGCGGCGAGATTGATCGCTGCTGCTGACTAACCCGATACCCGC 3661
DB |||||
QY 2407 TAGCGATGTACGCAATTCAGGCGCAGCTGAGCGCGCAGCGGTTGCCAGTGGCGCG 2466
DB |||||
QY 3662 AGCGTATGACAAAATTCAGGCGCATTTGCTGCGGCGGCGCACACAGCGCGCGCG 3721
DB |||||
QY 2467 ACATCGTCAGCAGGCGCTGTGCCAGCAGTCTGCTGGGAGGCGACTGATGCG 2526
DB |||||
QY 3722 ATATTGTTGAACAGGCGATGCG- GACCTGTGACCCAGTACTCAGTGGGCGAGGATTAIGA 3780
DB |||||
QY 2527 CAGCAATACGATGTGATTTTGGTGTGCTGAGCTGGGGAATGGCTTGATTGCGCTGCG 2586
DB |||||
QY 3781 ACCGCACTATGATCTCATTTCTGTGCTGCGGTCTGGCTAATGGCTTTATCGGCTCG 3840
DB |||||
QY 2587 TCTGCGTCAATTCAGCCACAACTGAATGCCCTGTTGCTGGAGAGCGATCGCGATCCGCG 2646
DB |||||
QY 3841 GCTTCAGCAACAGCATCCGATATCGGATCTTGTCTTATGAGCGGCTCTGAGCGGG 3900
DB |||||
QY 2647 AGGCAATCACTGCTGCTGTTTTCATCACGCAATCTCAGCGCGCAACACTTCGCTGGCT 2706
DB |||||
QY 3901 AGGGAACCACTGCTGCTTTCACGAAGAGGATTTAAACGCTGAATCAGCATCGCTGAT 3960
DB |||||
QY 2707 GCAACCGCTGATTTACCGTGGTGGTTCAGGTTATCAGGTGCGTTTCTGCTGCGCTGCGCG 2766
DB |||||
QY 3961 AGCGCGCTTGTGTCTCATCTGCGCGGACTACCAGGTTCTGTTTCCCCCAACGCGCTCG 4020
DB |||||
QY 2767 CAATCTGGACGGGATTAATTTGCTCATCGCATCAGGGGATTTTCCGCGCATTTTACGC 2826
DB |||||
QY 4021 CCATGTGAACAGTGGCTACTACTGCGTGACCTCCCGCAATTTCCGCGGATATCTCCGGA 4080
DB |||||
QY 2827 GCGCATGGGTGACGATCTGTGGACAAACACAGCGCTACACAGAGTAAACCCACGAGGT 2886
DB |||||
QY 4081 ACAGTTTGGACAAATTTATGGCTGCTATCCGCGGTTTCAGCGGTTTCATGCTGAATCGGT 4140
DB |||||

QY 2887 GAGCTGGCGGATGGCGCTGAACTTGTGTCGCAAGTGGTGAATGATGCTGCGCGCTGCA 2946
DB |||||
QY 4141 CCAGTTAGCGGATGGCGGATTAATCATGCCAGTACAGTACGATCGACGACGGGTTACAC 4200
DB |||||
QY 2947 GCGACGCCCAATCTGAGCTGGGTATCAGGTGTTTCTTGGACAAGAGTGGCAGCTGGC 3006
DB |||||
QY 4201 GCCTGATTTCTGCATACGCGTAGGATTCAGGCAATTTATCGGTGAGGAGTGGCACTGAG 4260
DB |||||
QY 3007 GCAGCCGACGGCTGCGACGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 3066
DB |||||
QY 4261 CGCGCGCATGTTTATTCGTCACCGATATCATGATGATGATGATGATGATGATGATGATGAT 4320
DB |||||
QY 3067 TTATCGTTTTGCTTACACGCTGCGCTCAGCGCGATTCGGCTATTTGATGAAGATACCCA 3126
DB |||||
QY 4321 CTACCGCTTTGTTTATACCTGCGCTTTCGCAACCGCACTGCTGATCGAAGACACACA 4380
DB |||||
QY 3127 TTAGCTTAAACAGCCCGCTGGCGGAGAACACCGCTGCTGACGACATTCGCGACTATGC 3186
DB |||||
QY 4381 CTACATTTGACAAGGCTAATCTTACGGCCGAAACGGGCGCTCAGAAACATTCGCGATTAATGC 4440
DB |||||
QY 3187 CAATCAGCAGGCTGGACGCTAGTACGCTGCTGCTGGAAGACACGCGCATATTTACCGAT 3246
DB |||||
QY 4441 TCGCGGACAGGTTGGCGGTTACAGAGTTCGCTGCGGGAAGAACAGGTTGCAITGGCCAT 4500
DB |||||
QY 3247 TACCTTGAAGCGGCAACATTCGATTCGTCGAACAGCAGCGCGGCCAAGGTCAGCGG 3306
DB |||||
QY 4501 TACGTTAACGGCGGATTAATCGTCAGTTTGGCAACAGCAAC- - -CGCAAGCTGTAGCGG 4557
DB |||||
QY 3307 CTTGCGCGCGGCTGTTTTCATGCGCACACCGGTTACTCCTTGGCGCTCGCGGTCGCT 3366
DB |||||
QY 4558 ATTACGCGCGGCTGTTTTCATCGCAACCGGCTACTCCCTACCGCTCGCGGTGGGCT 4617
DB |||||
QY 3367 AGCGGATTTGATGAGCAGCGCTGTTGGCCACCGATGCGCTCAGCTCAGCCAAACATATCGA 3426
DB |||||
QY 4618 GCGCGATCGTCTCAGCGCGCTGATGTTTACCTTCTCTCTCTCCTCCTCAGCAGCATTCG 4677
DB |||||
QY 3427 ACCTTTGCGCGTCAAGAGTGGCGGAAACAGCGATTTTTCGCTGCTGCTAAACCGCATGCT 3486
DB |||||
QY 4678 TCACTTTGCCAGCAACGTTGCGCAACAGGGGTTTTTCCGCAATGCTGAATCCGATGTT 4737
DB |||||
QY 3487 GTTTTGGCGGTAAGCGCGCAGCGCTGCGGCTGATGCAAGTGTATTTTACCAGCTGCA 3546
DB |||||
QY 4738 GTTTTGAAGCGGACCGCGCGAGTCAAGCTGCGGTGATGAGCGGTTTCTATGCTTACC 4797
DB |||||
QY 3547 TGCGGGTTAATTAGCGCTTTTACGCGGCAACTGCGCTGCGCGGATAAAAACCGGAT 3606
DB |||||
QY 4798 CGAGGATTTGATTTGCCGCTTTTATGCGGAAAACTCACCGTGACCGATCGGCTACGCAT 4857
DB |||||
QY 3607 TCTGTCGCGCAAGCGCGGTCGCCATCGGTGAAGCGCTGCGCGCTGTTGAATTCGT 3666
DB |||||
QY 4858 TCTGAGCGGCAAGCGCGCTTCCCGTTTCCGCGCATTTGAGGCAATTAAGCAGCTCA 4917
DB |||||
QY 3667 CGAACCGGGAAGAAAAATGAACGCACTTATGATTTGGCGCAGGCTTTGGCGGCTG 3726
DB |||||
QY 4918 TCGTTGAAGAGGACTACATGAACCACTACGTTAATTTGTCGGGCTTTTGGTGGCTG 4977
DB |||||
QY 3727 GCGTGGCGATTCGCTCAAGCGCGGCAATCAACACACTTACTCGAGAGCGCGAC 3786
DB |||||
QY 4978 GCACCTGGCAATTCGTTTACAGCGCGAGTATTCCTGTTTGTGCTGTTGAGCAGCGCAC 5037
DB |||||
QY 3787 AAACCGGCGGACGCGCTTATGTTTTCAGGACAGTGGCTTTTACCTTCGATGCGCGACCC 3846
DB |||||
QY 5038 AAGCGGGTGGCGGGCTTATGTTTATCAGGACAGGCTTTTACTTTTGTATGACGCGCT 5097
DB |||||
QY 3847 ACCTGATCACCAGTCCAGCGCCATCGAAGAGTGTTCACGCTGGCAGGAAAAATCGCTC 3906
DB |||||
QY 5098 ACCGTTATCACCGATCCCGAGCGGATTTGAAGAACTGTTTGTCTGCGCGGTAACAGCTT 5157
DB |||||
QY 3907 AGCGATTACGTCGAGTGAATGCGGTAAACCGCTTCTATCGCTGCTGCTGGGAAGATGCG 3966
DB |||||
QY 5158 AAGGATTACGTCGAGCTGTGCGCGTTCGCGGTTTTATCGCGCTGTGCTGGAGTCCGCG 5217
DB |||||
QY 3967 AAACAGCTTGAATTACGACAAATTAATCAGCGCGCTGCTGGAGCAGCAGATCGCCACGTTCAAT 4026
DB |||||

QY	6187	CACCGCATGATGCGGGTCCCATGTATTAAGCGCTTAAGATAGCCTTTCGCGGGAT	6246
Db	7429	TACAGCATGATGATACGGTGGCCATGTATACCCTTCAGGTAGCCTTTCGCGGGTAT	7488
QY	6247	ATACCGGAACGGCCAGCGTTGATGCACCGGCCATCGTGACCAATGAAGTAGAGCGCGC	6306
Db	7489	GTACCGGAACGGCCAGCGCTGGTATCCATCTCGTGTGGACCATTAATACAGTAAACC	7548
QY	6307	GTACGTCGTCAATTCGGGCACCAATCCACTGCAGCGGCACATCGCTTGCACACCGACATA	6366
Db	7549	ATAAGCGGTCAATGCTGTCACCAATCCACTGGAGCGGCCAGATTCTGTACTGCCGAAGTA	7608
QY	6367	AATCAGCACAAATCGCCAGTAGTACCGCAACACCCACCGCATAAAGATCGTTGAGCTCAACATT	6426
Db	7609	AATCAGGGCAATCGACAAATGGCGAATATACCGGCATAGAGATCGTTAACTTCAATGC	7668
QY	6427	ACCGCTGTCGGTTCATGGTGCAGACAGATGCCAGCCCCATCCCCAACCGTGCATGATGA	6486
Db	7669	GCCTTTTACGGGGTTCATGATGTGAAGATGCGAGCCCCAACCCACGCGTGCATGATGA	7728
QY	6487	TTTATGGACAGCCCGCTACGATTTCCATCATCACACCAGCGTGGCCAAACAGATATAGCAC	6546
Db	7729	TTTATGTGCCAGTCAGCAACCACTTCCATGCGGACCAACCGTGCACAAACAGATCAGGC	7788
QY	6547	GTTCCATAAACGAGCATTTGTTCTCCA	6574
Db	7789	ATTCCAATCCCAACATAATTTCTCAA	7816

RESULT 6
ADM98599
ID ADM98599 standard; DNA; 12753 BP.
XX
XX
AC ADM98599;
XX
XX
DT 01-JUL-2004 (first entry)
XX
XX
DE Geranylgeranyl pyrophosphate synthase DNA #19.
XX
XX
KW Geranylgeranyl pyrophosphate synthase; gene; ds; diterpene;
KW diterpene precursor; diterpene synthase; defence toxin;
KW valstatin defensive signal; pollinator attractant; photoprotectant.

XX	Pantoea agglomerans.
OS	
XX	US2004072323-A1.
PN	
XX	15-APR-2004.
PD	
XX	07-JAN-2002; 2002US-00041018.
XX	
XX	05-JAN-2001; 2001US-0259880P.
PR	
XX	(MATS/) MATSUDA S P T.
PA	(HART/) HART E A.
PA	
XX	Matsuda SPT, Hart EA;
PI	
XX	WPI; 2004-373921/35.
DR	
XX	New unicellular organisms comprising exogenous nucleic acids encoding a
PT	geranylgeranyl pyrophosphate and a diterpene synthase, useful for
PT	producing diterpenes and diterpene precursors.
PT	
ps	Claim 2: SEQ ID NO 19: 38bp; English.
ps	

XX The invention relates to a unicellular organism for producing a diterpene
CC or diterpene precursor comprising an exogenous nucleic acid sequence
CC encoding a geranylgeranyl pyrophosphate synthase under the control of a
CC promoter operable in the organism, and an exogenous nucleic acid sequence
CC encoding a diterpene synthase under the control of a promoter operable in
CC the organism. The invention also relates to methods of producing a
CC

diterpene or diterpene precursor and a method of isolating a diterpene synthase comprising growing several cells in the presence of a polyaromatic resin to make a cell/resin mixture, where at least one of the cells further comprises at least one isolated and purified nucleic acid sequence of a yeast expression library, and the expression of the nucleic acid sequence is regulated by an inducible promoter under conditions where the expression is induced, filtering the cell/resin mixture, extracting the cell/resin mixture with alcohol to produce an organic eluent and analysing the organic eluent by a screening method including chromatography and/or spectroscopy, to identify the nucleic acid sequence encoding the diterpene synthase. The unicellular microorganism is useful as a diterpene or diterpene precursor producing system. Diterpenes, in plants, serve as defence toxins, volatile defensive signals, pollinator attractants and photoprotectants. This sequence represents DNA encoding a geranylgeranyl pyrophosphate synthase polypeptide of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

XX	Sequence	12753 BP;	2489 A;	3880 C;	3911 G;	2473 T;	0 U;	0 Other;
QY	Query Match	30.5%;	Score	2135.2;	DB	12;	Length	12753;
QY	Best Local Similarity	64.4%;	Pred.	No. 0;				
QY	Matches 3263;	Conservative		0;	Mismatches	1765;	Indels	39;
QY								Gaps 4;
QY	1479	CTGTTTCGGGTGATCGACGATCTCGCGTCTGTCACCGATATGCTGTGTCGCGGAATGGCT	1538					
QY	5780	CTGCTGGGCTGATTGCGGCCATGTCATCCCTGACGATGTGCTGTGCACGACGATTGGCC	5839					
QY	1539	GCGGTACTGAAGCATTGAACATCGATGCGGTGATCGCGCGAATGAAAGCGCGCGGC	1598					
QY	5840	GCTATTCTACAGCGGCTGCGGTGGACGCGCTGATTGTGATGATGAGATGAGCGCCGCGA	5899					
QY	1599	GGATTGTCCTGAAGCGCTGCATCTGCGGTGTTGTTGCGTGGCTGCGCCTTGCCTGCGGTC	1658					
QY	5900	AGCTGTGTCGCGGCGCTGGACTACCATTTATCTCTATTGCTGCGCGTGTGCGGTC	5959					
QY	1659	AATCGTGAAGCGGGATTCGCTTGGCGGTGATGCCCTTCGCTTTTGCACAGGATGACAAA	1718					
QY	5960	AACCGGAGCGGGCTGTGCGCTGCGCGTGTGCGGTTTCACTACGCCGAGGTAAGAGA	6019					
QY	1719	GCGCTGAAACGTTTTTCAGGGCCAGCAGCGATATCTATGATCGCATCATGCGTCGTCA	1778					
QY	6020	GCCCTGCGCGGTTTTTCHAGTTCACGAAACGATCTACGATGCGTGTATGTTACCCG	6079					
QY	1779	GACGTGATCTCTAAACACGCGCGCGGCTTAAATTGACGGAAGCGGCGGATTAATCAG	1838					
QY	6080	CAGACGATCCTTGGGCCACGCCACGCGTTTGGTTTGCAGGAGCGCAGCGCTCTCAG	6139					
QY	1839	TGCTGTGCGGCTGGGACAAATCAGCCAGATGGTCGCGCTTTGATTTTCCAGTCTCAG	1898					
QY	6140	TGCTCTCGCGCTGGGCGAGATGAGCCAGTCCGTTTCCGCGCTCTGACTTCCACGCGG	6199					
QY	1899	CAACTGCCCGCTCTACTACGCGCTGGGCGCACTCCGCGCCCGTTCCTCTGCGCGG	1958					
QY	6200	GCGCTGCGGAATGTTTTTCTACTAGTGGGAGCACTGCGC-----TATCAGCGCGCG	6250					
QY	1959	CTCCATGCGCGCTGGCAGCGCTGCTGTCACCGGTGGTTTATGCTCTGCTGGGTACGCTG	2018					
QY	6251	CGCAGGTAGAACGCTGCCACGCGACGCGCCGCGATCTTTGCTCTGCTGGGCAACCTTC	6310					
QY	2019	CAAGGCATCGCTTCCGCGCTGTTTTCTGCATCTGGCGCAGCGCTGGCGCCAGCTGCGGCTA	2078					
QY	6311	CAGGCGCCACCGTCTACGCTGTTTTTCAAGATGCGCGCGCTGTGCCAGCTGGGGCGG	6370					
QY	2079	TGCTGTGTATCGGCCATTGTGGGGGATTAACGCGGAAACAGACGATCAGCTGGAGCTC	2138					
QY	6371	GAGGTGACCATTTGCCACTGCGATGCGCTGACGCGCCCGCCAGGCGCATCTCGCTCTAGCC	6430					
QY	2139	GCTGCGCGGCTGGGTGACGGATTTGTCGATCAGCGCGCAGCGCTTACAGCAGCGGCGAG	2198					
QY	6431	TGCGCGCGGACGGAGGTGGTTCAGCTTTGTTCGACCGCGCGCTACGTTGTCGAGGCTAAT	6490					

Db 8627 GAACGGCATGTGTAAGACTGTTTACCGATCTGGCGGGGAGATCGAACTCAACGCCCGGGT 8686
Qy 4416 GAGCGAGTGGAAACACAGCGGCAATCGCATAGCGCGCTTCAGTTAGAGGGCGGACGACG 4475
Db 8687 CGAAGACTGGTGGCGGATACCGGCTAAGCCAGGTCCGGCTGGCGGATGGTCGGAT 8746
Qy 4476 TTTCGATCCCGCGCTGTGGCTTCAATTCGCGACGTGTGCGATACCTACGCAAACTGCT 4535
Db 8747 TTTCGACACCGACGCGGTAGCTCGAAACGCTGACGTGTGAACCTATATAAAGCTGCT 8806
Qy 4536 TCGCACCATCCGCTGGCAATGAACAGTCGCGACATCGCTGAAGCTGAAGCGATGAGCAA 4595
Db 8807 CGGCAACCATCCGGTGGGCGAAGCGGGCGGCGAGCGCTGGAGCGCAAGAGCATGAGCAA 8866
Qy 4596 CTCGCTGTTTGTACTTATTTGGCTGAATCAGCCGATGAACAGCTCGCGCACACAC 4655
Db 8867 CTCGCTGTTTGTCTTACTTTCGGCTGNAACAGCTCATTTCCAGCTGGCGCACCATAC 8926
Qy 4656 CCTCTGTTTGGCCCGCTTATCTGTAGTGTGATCGATGAGATTTTCAACAGCAGCCAGCT 4715
Db 8927 CATCTGTTTGGTCCCGCTACCGGAGCTGATCGACGAGATCTTTACCGGCGCGCT 8986
Qy 4716 GCGACGATTTTCACTTTACTCGACGCGCTCGACGCGATCCGTCTGCGCAC 4775
Db 8987 GCGGATGACTTCTCGCTCTACTCTGCTGCTGCGCTGCGTGAACGATCCCTCGCGCC 9046
Qy 4776 GCCCGCTCGCGCAGCTTTATGTGTAGCGCGGTGCGCATCTCGGCACCGCTGACAT 4835
Db 9047 TCCCGCTGCGCAGCTTCTAGTGTCTGCGCCGCTGCGCATCTTGGCAACGCGCGT 9106
Qy 4836 CGACTGGCAACAGGAAGACCGGCTTGGCGCATCGAATTTTGGCTTATCTGGAGCAGCA 4895
Db 9107 GGACTGGGCGCAGGAGGGCCGAAGCTGGCGCACCGCATCTTTGACTACTTTGAAGAGCG 9166
Qy 4896 CTACATGCGGATTTACGTCAGCAATTAGTGACACAGAAATTTTACGCCGTTTGTAT 4955
Db 9167 CTATATGCGCGCTGCGTAGCGAGCTGGTGACCCAGCGGATCTTTACCCGCGCAGCT 9226
Qy 4956 TCGCACACGCTGCAATGCCATCAGCGCTCGCGGTTTTGCTGGAGCGGATTTGACGCA 5015
Db 9227 CCAAGACACGCTGGATCGCATCTGGATCGGCTTCTCCATCGAGCGCTGCTGACCCA 9286
Qy 5016 AAGCGCTGCTGCGCCCGCATACCGCGATCGGATTCAGCAATCTATCTGTTGGG 5075
Db 9287 AAGCGCTGCTGCGCCCGCACAACCGCGACGAGCATTTGCCAACCTCTACCTGTTGGG 9346
Qy 5076 TCGCGTACGATCCAGGCGCGGCTGCGCGGTGATCGGTTGCGGCCAAGCCACCGC 5135
Db 9347 CGCAGTACTCACCTTGGGGGGGCAATTCCTGGCGTAGTGGCTCGGCGAAGCCACCGC 9406
Qy 5136 CAGGCTGATGCTGGAGGATCGCGCCGAATGAATCGACAGCTTTTACTTGAGCAAGTAAG 5195
Db 9407 CAGCCTGATGTTAGGATC---TGCAATGAGCAACCGCGCTGCTTCAACACGCCACG 9463
Qy 5196 CAAACATGGGTGGGCTGAAGAGTTTCGCAACCGCCCGCAAGCTGTTGATCGACCG 5255
Db 9464 CAGACATGGCCAAACCGGCTCGAAAAGTTTTGCGACCGCTGCGAAGCTGTTCGACCGGCG 9523
Qy 5256 AGCGCGGAGCAGCAGCTGATGCTGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5315
Db 9524 ACCCGCGTAGCTGCTGATGCTCTACCTGCTGCGGCGCACTGCGATGAGCTATTGAC 9583
Qy 5316 GGGCAACCGTGGGCGAAGCGCGCACGCGATGCGCTCGAAGACGCGCAGGCGATGATG 5375
Db 9584 GACCAGACCGCTTCCGACGAGGCGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9643
Qy 5376 CAGCATCTGAATTAACCCCGCGCTCAGCGCGCGCAGCATGAGTGAACCGCGC 5435
Db 9644 GCCCGCTGCGCAGCTGACCTTGGCGGCTTTGAAGGGGCGGAGATGAGGATCCGGCC 9703
Qy 5436 TTTAGGCGGCTTCAGGAAGTGGCGATCAATTCACAGCTGCGCGCAACAACTGGCGTTTGTAT 5495
Db 9704 TTGCTGCTTTCAGGAGGTGGGCTGACCCAGGTTTACGCGCCCGATGGCGCTGCTGAT 9763

Qy 5496 CATCTGAAGGCTTTGCTATGATGCGCAACAAACATTAACGCGAGCTTCGATGACAG 5555
Db 9764 CACCTCAGCGCTTTGCGATGGACGTGGCTCAGACCCGCTATGTCACTTTTGAAGATACG 9823
Qy 5556 CTGCGTTACTGCTATCAGCTCGCGGCGTGGTGGTTTGTATGATGGCGCGTAAATGGGC 5615
Db 9824 CTGCGCTACTGCTATCAGCTGGCGGCGTGGTGGTTCTGATGATGGCCAGGTTGATGGC 9883
Qy 5616 GTGCGCAGCAAGCGGTCTGATCAGCGCTGCGATTTAGGACTTGGGCTTCACGCTCACT 5675
Db 9884 GTGCGGATGAGCGGCTGCTGATCGCGCTGCGATCTGGGCTGGCTTCACGCTGACG 9943
Qy 5676 AACATTGCGCGCATTTGTAGAAAGATGCGCAAAATGGTCTGCTGCTATCTGCCGCAATCC 5735
Db 9944 AATATCGCCCGGATATTAATGAGATCGGCTATTGACCGCTGCTATCTGCCCGCGAG 10003
Qy 5736 TGGCTCGATCAGGCGGATTAAGCGCGGATACGCTGATCGCACCGCAACATCGTGCAGCG 5795
Db 10004 TGGCTGAGGATGCGGCTGACCCCGGAGAACTATGCGCGCGGAGAAATCGGCGCGC 10063
Qy 5796 CTCGCTCAGCTGGCAGCGGCTTTAGTGGCGAGCGGAACCTTATTAATCACTCGCGCGCA 5855
Db 10064 CTGGCGGCTGGCGAGCGGCTTTATGATGCGCGAGCGCTACTACATCTCTCCCGAG 10123
Qy 5856 TCGGTTTACCGGTTTACCGCTGCGCTGGCGTGGGCGCATCGCTACGGCTCGCGGCTT 5915
Db 10124 GCGCGCTTACAGATGCGCGCGCTGCGCTGGGCGATCGCAACCGCGCGAGGCTC 10183
Qy 5916 TATCGCAAAATTTGCGCTCAAAATTTGACGACCGCGGTGTCACGCTGGGATTCACGGCAG 5975
Db 10184 TACCGGAGATCGGTTATTAAGTAAAGCGCGGAGGCGCGCTGGGATCGCGCGCAG 10243
Qy 5976 GCGACAGTAAGGTGAAAACTGCGCTGCTGTAAGGAGGCGAGGTTTGGCGATCACT 6035
Db 10244 CACACCAAGGTGAAAAATTTGCCATGCTGATGGCGCACCGCGGCGAGTTTTCGG 10303
Qy 6036 TCGGCTGCTGCTGCTGTAACCGGCTCGGCTGGTCTGTTGGCAGCGCTTCGTTGATTT 6095
Db 10304 GCGAAGCAGCAGGCTGACCGCGCTCGCGCGCTTTTGGCAGCGCTTCGTTGAGCG 10363
Qy 6096 TACGCTCGTACGCTGGCGCAGCGTGGCTTCAGCTTATTCAGCGGTGGCGCTGAGAGA 6155
Db 10364 GCGCGCATGAGCTTCAGCAGGATCGCTGTAGTGGCAGGCTTGGCGGCTTAAATAA 10423
Qy 6156 AACCAACGACACGCGCTTTCAGCGCGCGCACCGCATGATGATGCGGTGGCGCATGT 6215
Db 10424 AACCAAGAGACGCGCGCTTCCGCGCGCACCGCGTGGTGGCAGCGCTTCGTTGAGCG 10483
Qy 6216 ATAGCGCTTAAGATAGCTTTTGGCGGATATAGCGAAGCGCGCGGTTGATGACCA 6275
Db 10484 AGACCGCTTACGTTAGCTTGGCGGCGGATTCAGTGAAGGCGCGCTGATGACCA 10543
Qy 6276 GCGCATCTGCAACATGAAGTAGCGCGCGCTTACGCTGCTATTCGCGCACCAATCCACT 6335
Db 10544 GACCGCTGTCACAGAGTAGAGCGGCTATAGACCGCTCATGCGCGAGCAATCCACT 10603
Qy 6336 GCGCGGCGCATGCTTTGCAACCGCATATAATCAGCAATTCGCCAGTACCGCAACA 6395
Db 10604 GCGGGGCGAAACCGCGCGCTGCGCAACGCGCAATCAGCGCGATAGCCACCGCGCAACA 10663
Qy 6396 CCACCGCATTAAGATCGTTGAGCTCAAACTTACCGCTGCGGTTTCATGTTGCGGACAGAT 6455
Db 10664 CCACCGCAAGAGATCGTTTACCTCAAAATACGCCCTTGGCGGGGATGTTGTTGACTCAT 10723
Qy 6456 GCGACCGCATCCCAACCGCTGATGATGATTTATGCGACAGCGCGCTACGATTTCCA 6515
Db 10724 GCGAGCGCATCCCGCGCGCTGATTAATGATGCGGTGGGTAACCGCGGATGCCCTCCA 10783
Qy 6516 TCACCAACAGCGTTGCCAACAAGATAA 6542
Db 10784 TCGCAATAACGCTCAAGATGACGATTA 10810

QY 3305 GGCCTGCGCGCGGCTGTTTCATGCCACACCGGTTACTCCTTGCCGTGCGCGTGGCG 3364
DB 908 GGATTACGTCGGGCTGTTCCATCTTACCACCGCTATTCTACGCGCTGCGGTTGCC 967
QY 3365 CTAGCGGAGTTGGTAGCAGCGCTGTTGCCACCGATGCCCTCACGCTCAGCCAACTATC 3424
DB 968 GTGGCGGACCGCTGAGTGCACTTGATGTCTTTACGTGCGGCTCAATTCACCATGCCAAT 1027
QY 3425 GAAAGCTTTGCCCGTCAGCAGTGGCGGCAACAGCGATTTTCCGCTCTGTAACCGCATG 3484
DB 1028 AGCAATTTGCCCGGAGCGCTGGCAGCGAGCGCTTTTCCGATGCTGATCGCATG 1087
QY 3485 CTGTTTTGGCGGTAAGCCGAGCAGCGCTGCGCGTGAAGCAACGTTTTTACCGGCTC 3544
DB 1088 CTGTTTTAGCCGCGCCCGCGATTCACGCTGGCGGTTATGACGCTTTTATGCTTTA 1147
QY 3545 GATCGCGGTTAATTTAGCGCTTTTACGCGGGGAACTGCGGCTCGCGATAAACCGCGG 3604
DB 1148 CCTGAAGATTTAATTTGCCGCTTTTATGCGGGAACCTCACGCTGACCGTACGCT 1207
QY 3605 ATTCTGCGGCAAGCCGCGCTGCGCATTCGGTGAAGCGCTGCGCGCTGTTGAATTC 3664
DB 1208 ATTCTGCGGCAAGCGGCTGCTCGGTATTAGCAGCATTCGAAGCAATTAAGCACT 1267
QY 3665 GTCGAACCAAGGAAGAAATGAACGCACTTATGTGATTGGCGCAGGCTTTTGGCGCC 3724
DB 1268 CATCGTTAAGAGCGACTACATGAACCAACTACGTTAATTTGTCAGGCTTCGGTGGCC 1327
QY 3725 TGGCGCTGCGATTGCGCTGGAAGCGGCGGATACCAACACTTACTCGAGCAGCGG 3784
DB 1328 TGGCACTGGCAATTCGCTACAACTGCGGGATCCCGCTTACTGCTTGAACACGCTG 1387
QY 3785 ACAAAACCGGCGGCGCGCTATGTTGTTGAGCAGTGGCTTTTACCTTCGATGCGGAC 3844
DB 1388 ATAAACCGGCGGCTGCGGCTTAATGCTACGAGATCAGGGGTTTACTTGTATGAGGCC 1447
QY 3845 CCACGGTATACCGGATCCAGCGCATCGAAGAGTTGTTACGCTGGCAGGAAATCGC 3904
DB 1448 CGACGGTTATACCGGATCCAGTCCATTTGAAGAACTGTTGCACTGGCAGGAAACAGT 1507
QY 3905 TCAGCGATTACGTCGAGCTGATGCGGTAAACCGCTTCTATCGCTGCTGCTGGAGATG 3964
DB 1508 TAAAGAGTATTCGAACCTGCTGCGGTTACGCGCTTTTACCGCTGCTGTTGGAGTCA 1567
QY 3965 GCAACAGCTTGATTACGACAATAATCAGCCGCTGCTGGAGCAGCAGATCGCACGTTCA 4024
DB 1568 GGAAGTCTTTAATTTAGATACGATCAACCCGCTCGAAGCGAGTTTACGAGTTA 1627
QY 4025 ATCCGCAAGATGTAGAGGCTATCGTCAATTTCTTGCCTATTTCAGTGAAGTATTAGAG 4084
DB 1628 ATCCCGCGATGTCGAAGGTTATCGTCAGTTTCTGAGCTATTTCACGCGGCTGTTTAAAG 1687
QY 4085 AGGTTATCTGAACTCGGCAAGTGGCGCTTCTGCGAGTGGTGACATGCTGCGGCTG 4144
DB 1688 AAGGCTATCTAAAGCTCGGTGCTGCTTTTATTCGTTTCAGAGACATGCTTCGCGCG 1747
QY 4145 CCGCGAGTTGGGACGCTGCAAGCATGCGCAGCGCTTACAGCATGTTGGCGGAAATTTA 4204
DB 1748 CACCTCACTGGGAACTGCGAGCATGGAAGCGTTTACAGTAAGTTGCGAGTTACA 1807
QY 4205 TTCAAGCATCATCTGCGTCAAGGCTTTTCTTCCACTCATTTGCTGTTGGCGGTAATC 4264
DB 1808 TCGAAGATGAACATCTGCGCCAGGCGTTTCTTCCACTCGCTGTTGTTGGCGGCAATC 1867
QY 4265 CTTTTCACGCTCATCTATCTAATTTATGCTGCTGAGCGGTGATGCGGCTGT 4324
DB 1868 CCTTGCACCTCATCTCAATTTATGATACACGCTGAGCGGTGATGGGCGCT 1927
QY 4325 GCTTTCGCGCGCGGCAACCGCGCTGTCAGGCAATGCGGCACTGTTTTCAGGACT 4384
DB 1928 GCTTTCGCGCTGGCGGCAACCGGCAATTTAGTTACGGGATGATAAGCTGTTTTCAGGATC 1987
QY 4385 TGGCGGCGAGCTGTTTACTGAATGCGGAAGTGAAGCCAGCTGGAAACCGCGGCAATCGCA 4444

DB 1988 TGGTGGCGAAGTCGTGTTTAAACCGCAGAGTCAGCCATATGGAACACGACGAAACAGA 2047
QY 4445 TTAGCGGCTTTCAGTTAGAGGCGGACGCTTCGATGCGCGCTGTGCTTCCCAATG 4504
DB 2048 TTGAAGCGCTGCATTTAGAGGACGCTCGAGGTTCTCGACGAAGCGCTGCGCTCAATG 2107
QY 4505 CCGAGCTGGTCATACCTTACGACAAATGCTTCCGACCACTCCGCTGGCAATGAACGTC 4564
DB 2108 CAGATGTTTCATACCTATCGCACCTGTTTAAAGCCAGCACCTGCGCGGTAAAGCAGT 2167
QY 4565 CGACATCGCTGAAGGTTAAGCGCATGAGCAACTGCTGTTTGTACTCTATTTTGGCCTGA 4624
DB 2168 CCAACAACTCGACACTAAGCGCATGAGTAACCTCTGTTTGTCTCTATTTTGGTTGA 2227
QY 4625 ATCAGCGCATGAACAGCTCGCGCACCAACCGCTCTGTTTGGCCGCTTATCGTAGT 4684
DB 2228 ATCACCATCATGATCAGCTCGGCATCACCGTTTGTTCGCGCGGTTTACCGGAGC 2287
QY 4685 TGATCGATGAGATTTTCAACAGCAGCAGCTGGCAGACGATTTTTCATCTTTACCTGCA 4744
DB 2288 TGATTGACGAATTTTAAATCATGATGGCTTCGACAGGACTTCTCACTTATCTGCACG 2347
QY 4745 CGCCTGCGAGCAGCATCCGTCGCTGCGA CCGCCCGCTGCGGAGCTTTTATGTGTAG 4804
DB 2348 CGCCTGCTGTCAAGATTCGTCACCTGCGCGCTGAAGGTTGCGGAGTTTACTATGTGTGG 2407
QY 4805 CGCGGTCGCGCATCTCGGCACCGCTGACATCGACTGGAACAGAGGAGCGCGCTTGC 4864
DB 2408 CGCGGTCGCGCATTTAGGACCGCGAACCTCGACTGGAAGTTGAGGGGCAAACTAC 2467
QY 4865 CGCATCGAATTTTGTCTTATCTGGAGCAGCACTACATCGCGGATTAAGTCAGCAATPAG 4924
DB 2468 CGCAGCTGTTTGGCTTACCTTGAGCAGCATTACATGCTGCTTACGAGTCACTGCTG 2527
QY 4925 TGACACAGAAATGTTTACGCGCTTGAATTTTCGCA CAGCTGCTGATGCCATCACGCT 4984
DB 2528 TCACGCAACGAGTGTTCACGCGTTGATTTTCGCA CAGCTTAAATGCTTATCATGCT 2587
QY 4985 CGCGGTTTTCGCTGAGCGGATTTTGAACGAAAGCGCTGCTCGCCCGCATTAACCGCG 5044
DB 2588 CAGCTTTCGTTGGAGCGCTTCTTACCGAGAGGCTGTTTCGCGCGCATTAACCGCG 2647
QY 5045 ATGCGCATCAGCAATCTCTATCTGTTGGTGCGGTA CCGATCCAGCGCGGCGTGC 5104
DB 2648 ATAAACCATTTACTAATCTCTACCTGTCGCGCAGG CACGATCCCGCGCAGGCAATC 2707
QY 5105 CCGGCTGATCGGTTTCGCGCAAGGCCACCGCAGCTGATGCTGGAGGATCGCGCGAAT 5164
DB 2708 CTGGCTCATCGCTCGGCAAAAGCGACAGGTTTGTATGCTGGAGATC---TGATTT 2764
QY 5165 GAATCGACAGCTTTTACTTGAAGCAAGTAAACGAAACCATGCGGTGGGCTCGAAGAGTTT 5224
DB 2765 GAATAATCCGTCGTTACTCAATCATCGGTGCAACGATGTCAGTTGCTCGAAGTTT 2824
QY 5225 CGCCACCGCGCAGCTGTTTGTATGACACCGCGCGCAGCAGCTGATGCTGTATGC 5284
DB 2825 TGCGACGCTTCAAGTTTATTTGATGCAAAACCGCGCGCAGCTGATGCTCTACGC 2884
QY 5285 GTGTTGCTCATCTGCGATGATGATGATGGGCAACGCTGGCGGAAGCGGCAACCA 5344
DB 2885 CTGGTCGCGCATGTTGACGATGTTTGAAGATCAGACGCTGGGCTTTCAGGCGCGGCA 2944
QY 5345 GCATCGCTCGAAGACGCGCAGGCACTGATGCAATCTGCAATCTGCAAAATGGAACCGCGCGG 5404
DB 2945 GCCTGCTTACAAAGCCCGCAACACGCTGATGCAACTTGAGATGAACCGCGCAGG 3004
QY 5405 CTACAGCGCGCAGCATGAGTAACCGCGCTTTAGGGCGTTTACGAGTGGCGATCAT 5464
DB 3005 CTATCAGGATCGCAGATGCAACCGCGCTTTTTCGCGCTTTTCAGGAAGTGGCTATGC 3064
QY 5465 TCACAGCTGCGGCAACACTCGGCTTTGATCATCTGGAAGCTTCGCTATGATGTCAGC 5524

Db 3065 TCATGATATCGCCCGCTTACGCGTTGATCATCTGGAGGCTTCGCCATGATGTACG 3124
QY 5525 CAAACGAACATTACGAGAGCTTCATGACACGCTGCGTTACTGCTATCATCGTCGCGGCGT 5584
Db 3125 CGAAGCGCAATACAGCAACTGATGATACGCTGCGCTATTGCTATCAGTTGACAGGCGT 3184
QY 5585 GGTGCGTTGATGATGCGCGGCTTAATGCGGCTGCGGACGAGCGGTGCTGATCAGCG 5644
Db 3185 TGTGCGCTTGATGATGCGCAATCATGCGCGTGGGATACGCCACGCTGGACCGCGC 3244
QY 5645 CTGCGATTTAGACTGGCGCTCCAGCTCACTAACATTGCGCGCATATTGTAGAGATGC 5704
Db 3245 CTGTGACCTTGGCTGGCATTTGATTTGCTGACCAATATTGCTGCGGATATTGTGACGATGC 3304
QY 5705 CGAAATGCTGCTGCTATCTCGCGCAATCTGCGGCTGATCAGGCGGGATTACGCCCGA 5764
Db 3305 GCATCGGCGCGCTGTTATCTCGCGCAAGCTGGCTGGAGCATGAAGGTCTGAACAAAGA 3364
QY 5765 TAGCTGATGACCGCAACATGCTGACGCGCTGCGCTCACTGGCAGCGCGTTTGTAGTC 5824
Db 3365 GAATTATGCGGCACCTGAAACCGTACGCGCTGAGCGGTATCGCCCGCTGTTGGTGCA 3424
QY 5825 GGAGCGGAACCTTATTACTCTCGCGCGATCCGCTTTACCGGGTTTACCGCTCGCGTC 5884
Db 3425 GGAAGCAGAACTTACTATTGCTGCCACAGCGGCTGGCAGGGTTGCCCTCGGTC 3484
QY 5885 GCGGTGGCGCATGCTGCTACGCGTGGCGGCTTTATCGGAAATGCGGTCAAAAGTTCAACA 5944
Db 3485 CGCCTGGGCAATCGCTACGCGCAAGCAGGTTTACCGGAAATAGGTGTCAAAAGTTGAACA 3544
QY 5945 CGCGGTGTCAGCGCTGGATTACGCGAGCGCACCACTAAGGTGAATAACTGCGCT 6004
Db 3545 GCGCGGTGACGAGCTGGGATGACGCGCAGTCAACAGCACCGCGGAAATAAATTAACGCT 3604
QY 6005 GCTGTGAAAGGCGAGGTTTGGCGATCACTTTCGCGTGTGCTCGTCTGAAACCGCGTC 6064
Db 3605 GTGCTGGCGGCTCTGCTCAGCGCTTACTTCCGGATGCGGCTCATCTCCCGGCC 3664
QY 6065 GCGTGTGCTGTGCGAGCTCTGCTGTTATTTACGTCGCTGACGCTGGCGCAGCGTGGCT 6124
Db 3665 TCGCATCTCTGCGAGCGCGC-----TCTAGCGCATGCTTTTCCGGAGCGTGGC 3718
QY 6125 TCGAGTTATTACGCGTGGCGGTAGAGAAACCAACGACACGCGAGCTTCAGCGCCG 6184
Db 3719 TGAAGTTTTCAGCGGCGCGCATAGAGAGAACCAACCAACCACTTTCTTTGCCC 3778
QY 6185 CGCACGCGATGATGATGCGGTGCGCATGTATAGCGCTTAAGATAGCTTTTGGCGGG 6244
Db 3779 CTGACGCGTGTATGATACGCTGGCGCATATACACCGTTTGGAGTGGCTTGTGTA 3838
QY 6245 ATATAGCGGAACGCGAGCTTGTATGACACCGCGCATCGTACCATGAAGTAGAGCGG 6304
Db 3839 ATATAGCGGAATGGCGCAAGTTGATGACACCGCGCTGTCACCATAAATAGAGTAAT 3898
QY 6305 CCGTAGCTGCTATTTCGCGCAACATCCACTGCGGCGCATGCTTGCACCGCA 6364
Db 3899 CCATAGCGCTCATCTCGCGCAATCCACTGAGCGGCCACATCTCTGTACTGCGCCAGA 3958
QY 6365 TAAATCAGCAACATCCAGTACCGCAACACACCGCAATAAGATGTTGAGCTCAAC 6424
Db 3959 TAAATCAGCAGATCGATATGACGCAAAACACCGCATAAAGATCGTTAATCTCAAC 4018
QY 6425 TTACCGCTGTGCGGTTCATGTCGACAGATGCCAGCCCATCCCAACCGTGCATGATG 6484
Db 4019 GCACCTTTTACGCGGTTTATGATGTGAAGATGCCATCCCAACCCAGCGGTGATG 4078
QY 6485 TATTATGCGACAGCGCGCTAGATTTCATCACCACCGTTCGCAAGAGTAAGC 6544
Db 4079 TATTGTGTGCGAGTGCAGCAATCACTTCATGCGCAATCACCGTAAACGAAACGATCAG 4138
QY 6545 AGCTTCATAACACAGAGCATTTGCTGCTCATTTTGTGGAAGAGGAAGTA 6593
Db 4139 GCATTCGAATCAACAAATAATTTCTCGGTAGACAGCTGTGGACGA 4187

RESULT 8
ABZ69177

ID ABZ69177 standard; DNA; 7494 BP.

XX AC ABZ69177;

XX DT 11-AUG-2003 (first entry)

XX Vector construct p8Scaro14 coding sequence.

XX Biosynthesis; enzyme; antibacterial; protozoacide; herbicide; isoprenoid;
KW inhibitor; immunostimulant; gene; ds.

XX OS Unidentified.

XX OS Synthetic.

XX FN DE10201458-A1.

XX PD 17-OCT-2002.

XX PF 16-JAN-2002; 2002DE-01001458.

XX PR 11-APR-2001; 2001DE-01018166.

XX PR 22-JUN-2001; 2001DE-01030236.

XX PR 09-NOV-2001; 2001DE-01055084.

XX PA (BACH/) BACHER A.

XX PA (ROHD/) ROHDICH F.

XX PI Adam P, Amsingler S, Bacher A, Eisenreich W, Hecht S, Rohdich F;

XX DR WPI; 2003-185556/19.

XX PT New proteins involved in isoprenoid biosynthesis, useful in screening for

XX PT inhibitors, also new intermediates, potential therapeutic agents, nucleic

XX PT acids and antibodies.

XX PS Disclosure; Page 45-47; 78pp; German.

XX CC The present invention relates to proteins capable of the enzymatic

XX CC conversion of 2C-methyl-D-erythritol-2,4-cyclodiphosphate to 1-hydroxy-2-

XX CC methyl-2-butenyl-4-diphosphate. The proteins are useful in screening a

XX CC chemical library for inhibitors of isoprenoid biosynthesis. Cells and

XX CC organisms that express them are useful for in vivo/in vitro production of

XX CC isotopically labeled biosynthetic intermediates and products of the

XX CC mevalonate-independent pathway of isoprenoid synthesis. Such

XX CC intermediates and products are used to screen for genes, enzymes or

XX CC inhibitors of isoprenoid/terpenoid biosynthesis, potentially useful as

XX CC antibiotics, antimetabolites and herbicides and as immunostimulators,

XX CC particularly for treating bacterial infection. Antibodies against the

XX CC intermediates are useful for detecting pathogens, in body fluids. The

XX CC present sequence is a coding sequence used in the exemplification of the

XX CC invention

XX SQ Sequence 7494 BP; 1722 A; 1935 C; 2026 G; 1811 T; 0 U; 0 Other;

Query Match

Best Local Similarity 23.1%; Score 1618.8; DB 8; Length 7494;

Matches 2375; Conservative 0; Mismatches 1217; Indels 6; Gaps 2;

QY 2490 CAGCAGCAAGTCGTGCTGCGGAGCGGACCTGATCGCAGCAACATACGATGATGTTGG 2549

Db 3213 CACCGTTCAGCTGTCTTAAAGTGGAGCGGCTATCAACCGCATTTATGATCTGTTCTCG 3272

QY 2550 TCGGTGCTGGAGTGGCGAATGGCTTGATGCGTCTGCTGCTGCTCAATTGAGCCACAC 2609

Db 3273 TGGGGGCTGGAGTTCGGAATGCGCTTATCGCCCTCGCTTTTCAGCAGCAGCAACCTGATA 3332

QY 2610 TGAATGCTGTGTTGCTGGAGAGCGATGCGCATCCGCGAGGCAATCATACCTGCTGTTTC 2669

Db 3333 TCGGTATTTTGTCTTATCGACCGCCGCCAGCGGGGGGATCATACGTGTTCTTTC 3392

5550 GAACTCGACTGGAACGGTTGAGGGGCAAAACATACGGGACCGTATTTTTCGCTACCTTGA 5609
4890 GCAGCACTACATGCGCGGATTAAGTACAGCAATAGTAGACACACAGAAATGTTTACCGCGTT 4949
5610 GCAGCATTACATGCTGGCTTACGGAGTACAGTGGTTCAGCAGCCGGAATGTTTACGCGGT 5669
4950 TGAATTTGCGGACACGCTGCATGCCATACAGGCTCGGGTTCGCTGGAGCCGATTTT 5009
5670 TGAATTTGCGGACCAAGCTTAATGCTTATCATGGCTTCTGAGAGCCGCTTCT 5729
5010 GACGCAAGCGCTGTTCCGCGCGCATAAACCGGATCCGATATACAGCAATCTCTATCT 5069
5730 TACCCAGAGCGCTGTTTCGCGCGCATAAACCGGATTAACCAATTAATTAATCTTACTCT 5789
5070 GGTGGGTGCGGTAGCATCCAGGCGCGGCGTGCCTCGGCTGATCGGTTCCGCCCAAGGC 5129
5790 GGTGGCGGAGCAGCATCCCGCGGAGGCAATCTCTGGCGTATCTGGCTCGGCAAGC 5849
5130 CACCGGAGCTGATGCTGGAGATGCGCCGGAATGAATCGACAGCTTTTACTTGAGCA 5189
5850 GACAGCAGTTTGTGCTGGAGATC--TGATTTGAATAATCCGTCGTTACTCAATCAT 5906
5190 GTAACGCAACCATGCGGTGGCTCGAAGAGTTTCGCCACCGCGCAAGCTGTTTGTAT 5249
5907 GCGGTGCAACGATGCGTGGCTGGAAGTTTTCGACAGCTTCAAGTTATTTGAT 5966
5250 GACCGAGCGCGGAGCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 5309
5967 GCAAAACCGCGGAGCGTACTGATGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 6026
5310 ATTGATGGCAACCGTGGGCGAAGCGGCGACGCGATGCGCTGCAAGACGCGGAGCA 5369
6027 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6086
5370 CTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5429
6087 CTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6146
5430 CCGGCGTGGTGGGCTTTCAGGAGTGGGATCATTTACAGCTGCGGCAACAACTGGCG 5489
6147 CCGGCGTGGTGGGCTTTCAGGAGTGGGATCATTTACAGCTGCGGCAACAACTGGCG 6206
5490 TTTGATCATCTGGAAGCTTTCGATGATGATGATGATGATGATGATGATGATGATGAT 5549
6207 TTTGATCATCTGGAAGCTTTCGATGATGATGATGATGATGATGATGATGATGATGAT 6266
5550 GACAGCTGCTTACTGCTATCAGCTCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5609
6267 GATACGCTGCTTACTGCTATCAGCTCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6326
5610 ATGGGCGTGGGCGGAGCGGTGCTGATCAGCTGCGATTTAGGACTGCGCTTCCAG 5669
6327 ATGGGCGTGGGCGGAGTAACCGACGCTGGACCGGCTGCTGATGCTGCTGCTGCTGCTG 6386
5670 CTCATACTGCGCGGACATTTGAGAAATGCGGAAATGCTGCTGCTGCTGCTGCTGCTG 5729
6387 TTGACCAATTTGCTGCGGATTTGAGAGTATGCGATGCGGCTGCTGCTGCTGCTGCTG 6446
5730 CAATCTGCTGATCAGCGGATTTACGCGGATACGCTGATGCGACCGCAATCATGCT 5789
6447 GCAAGCTGCTGAGCATGAAGTCTGAACAAAGAGATTTATGCGGACCTGAAACCGT 6506
5790 GACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5849
6507 CAGGCGCTGAGCGGATTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6566
5850 GCGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5909
6567 GCGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6626
5910 GCGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5969

6267 CAGGTTTACCGGAAATAGGTGTCAAGTTGAACAGCGGCTGACAGCCCTGGGATCAG 6686
5970 CGGACGCAACAGTAAGGTGAAACCTGGCGCTGCTGGTGAAGGCGAGGTTGGCG 6029
6687 CGGAGTCAACAGCACACCGCCGAAATTAACGCTGCTGCTGGCGCTCTGGTCAGGC 6746
6030 ATCACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6087
6747 CTACTTCCGATCGGCTCATCTCCCGCCCTGCGCATCTCTGGCAGCGCGCG 6804

RESULT 9
ABZ69178
ID ABZ69178 standard; DNA; 8547 BP.
XX AC ABZ69178;
XX AC
DT 11-AUG-2003 (first entry)
XX Vector construct pACYCcarol4 coding sequence.
DE Biosynthesis; enzyme; antibacterial; protozoacide; herbicide; isoprenoid;
KW inhibitor; immunostimulant; gene; ds.
XX Unidentified.
OS Synthetic.
XX DE10201458-A1.
XX 17-OCT-2002.
XX 16-JAN-2002; 2002DE-01001458.
XX 11-APR-2001; 2001DE-01018166.
PR 22-JUN-2001; 2001DE-01030236.
PR 09-NOV-2001; 2001DE-01055084.
XX (BACH/) BACHER A.
PA (ROHD/) ROHDICH F.
XX Adam P, Amsingler S, Bacher A, Eisenreich W, Hecht S, Rohdich F;
WPI; 2003-185556/19.
XX
New proteins involved in isoprenoid biosynthesis, useful in screening for
inhibitors, also new intermediates, potential therapeutic agents, nucleic
acids and antibodies.
Example 7; Page 48-50; 78pp; German.
XX
The present invention relates to proteins capable of the enzymatic
conversion of 2C-methyl-D-erythritol-2,4-cyclodiphosphate to 1-hydroxy-2-
methyl-2-butenyl-4-diphosphate. The proteins are useful in screening a
chemical library for inhibitors of isoprenoid biosynthesis. Cells and
organisms that express them are useful for in vivo/in vitro production of
isotopically labeled biosynthetic intermediates and products of the
mevalonate-independent pathway of isoprenoid synthesis. Such
intermediates and products are used to screen for genes, enzymes or
inhibitors of isoprenoid/terpenoid biosynthesis, potentially useful as
antibiotics, antimetabolites and herbicides and as immunostimulators.
CC particularly for treating bacterial infection. Antibodies against the
intermediates are useful for detecting pathogens, in body fluids. The
present sequence is a coding sequence used in the exemplification of the
invention
XX
SQ Sequence 8547 BP; 1884 A; 2296 C; 2279 G; 2088 T; 0 U; 0 Other;
Query Match 23.1%; Score 1618.8; DB 8; Length 8547;
Best Local Similarity 66.0%; Pred. No. 0;
Matches 2375; Conservative 0; Mismatches 1217; Indels 6; Gaps 2;
QY 2490 CAGCAGCAAGTCGTCCTGGCGAGCGGACCTGATCGGACCGCAATACGATGATTGG 2549

Db 2845 CACCGGTGAGCTGCTTTAAGTGGAGCGGCTATGCAACCGCAATATGATCTGATTTCTCG 2904
Qy 2550 TCGGTCTGGACTGGCGAATGGCTTGATTTGGCTGCGTCTGGTCAATTTGAGCCACAAC 2609
Db 2905 TGGGGGCTGGACTCGCGAATGGCTTATCGCCCTGCGTCTTCAGCAGCAGCAACCTGATA 2964
Qy 2610 TGAATGCTGTTGCTGAGAGCGCATGGCATCCGGCAGGCAATCATACCTGGTCTGTTTC 2669
Db 2965 TGGGTATTTTGTCTATCAGCGCGCACCCAGGCGGGGGAATCATACGTGCTGATTTTC 3024
Qy 2670 ATCAGCGCATCTCAGCGCGCAACAACTTGGCTGCTGCAACCGCTGATTAACGGTGGTT 2729
Db 3025 ACCAGATGATTTGACTGAGAGCCCAACATCGTTGATAGCTCCGCTGGTGGTTCATCACT 3084
Qy 2730 GGTGAGTTTACAGGTGGTGTTCCTGCTGCGCGCAACTGAGCGGGGATTAATGTT 2789
Db 3085 GGGCGGACTATCAGGTACGCTTTCCACAGCGCGTCTGTAAGCTGAACAGCGGCTACTTTT 3144
Qy 2790 CCATCGCATCAGCGGATTTTGGCCGCCCATCTTTACGCGCGATGGGTGACGATCTGTGA 2849
Db 3145 GTATTACTTCTCAGCGTTTTCGCTGAGGTTTACAGCGACAGTTTGGCCCGCACTTGTGA 3204
Qy 2850 CAAACACAGCCGTAACAAGGTAAACCCACGAGGTGAGCTGCGGATGCGCGTGAAC 2909
Db 3205 TGGATACCGCGTCCAGAGGTTAATGCGGAATCTGTTCCGTTGAAAGGGTCAAGTTA 3264
Qy 2910 TTGCTGCGCAAGTGTGATGATGTCGCGGCTGACGCGACGACATCTGACGCTGG 2969
Db 3265 TCGGTGCGCGCGGTGATTTGACGGCGGGGTTATGCGGCAATTCAGCACTGACGCTGG 3324
Qy 2970 GTTATCAGGTGTTTCTTGACAAAGTGGCAGCTGGCGCAGCGCGCTTGACGAGC 3029
Db 3325 GCTTCCAGCGTTTATGSCAGGAATGSCGATTGAGCCACCGCATGGTTTATCGTCTC 3384
Qy 3030 CGATCTGATGATGACCGCTGATCAGCAAGCGGTTATCGTTTGTCTACAGCTGC 3089
Db 3385 CCATTCATGATGACCGCTGATCAGCAAAATGGTTATCGCTTCGTGTACAGCCTGC 3444
Qy 3090 CGCTCAGCGCGCATCGGCTATTGATTGAAGATACCCATTACGTTAAACAGCCCGCGTGG 3149
Db 3445 CGCTCTCGCGACGAGTTGTTAATTGAAGACAGCACTAATTAATGATGAGCAATTAG 3504
Qy 3150 CGGAGAACCCGCTCGTACGACATGCGCACTATGCCAATCAGCAAGCTGAGCGTGA 3209
Db 3505 ATTCCTGAATGCGCGGCAAAATATTGCGACTATGCGCGCAACAGGGTTGGCAGCTTC 3564
Qy 3210 GTACGCTGCTGCTGAGAGCAGGCAATTAACGATTAACCTGAGCGGCAACATCGATC 3269
Db 3565 AGACACTGCTGCGAGAAACAGGGCGCTTACCCCAATTAATCTGTCGGGCAATGCCAGC 3624
Qy 3270 GATTCTGGCAACAGCAGCGCGCCAGCGTGCAGCGGCTGCGCGCGGGCTGTTTCAATG 3329
Db 3625 CATCTGG--CAGCAGCGCCCTGCGCTGATGATGATTAATGCGCGCTGTTCCATC 3681
Qy 3330 CCACCAACCGGTTACTCTTTCGCTGCGCTGCGCTGAGGATTTGATGAGCGGCTGT 3389
Db 3682 CTACCAACCGGCTATTCACTGCGCTGCGGTTGCGGTCGCGCAACCGCTGAGTGCATTG 3741
Qy 3390 TGGCCACCATGCGCTCAGCTCAGCAACATATCAAGCTTTCGCGCTGAGTGCATTG 3449
Db 3742 ATGTCTTTTACGTGCGGCTCAATTCACATGCGCAATTAACGATTTTGGCCCGGCGCTGCG 3801
Qy 3450 GCGAACAGCGATTTTTCGCTGCTAAACCGCATGCTGTTTTCGCGGTAAGCGCGCAGC 3509
Db 3802 AGCAGCAGGGCTTTTTCGCAATGCTGAATCGCATGCTGTTTTCGCGGACCGCGGATT 3861
Qy 3510 AGCGCTGGCGGCTGATGCAAGGTTTTCAGCGGCTCGATGCGGGTTAATAGCGGCTTTT 3569
Db 3862 CAGCGTGGCGGTTATGACGCGTTTATGTTTACCTGAAGATTTAATTTGCGCGGTTT 3921
Qy 3570 ACGCGGCAACTGCGCTGCGGATAAACCGGATTTCTGTCGCGCAAGCGCGGCTGC 3629
Db 3922 ATCGGGGAAACACTCAGCTGACCGCTGCGGCTACGTTATCTGAGCGGCAAGCGCCCTGTT 3981

Qy 3630 CCATCGGTGAAGCGCTGCGCGCTGTTGAATTTCTGTAACAGGAGGAGGAGGAGGAGGAGG 3689
Db 3982 CGGTATTAGCAGCATTCGAAGCCATATGACGACTCATCGTTAAAGAGCGACTACATGAA 4041
Qy 3690 ACGCACTTATGTGATGTGGCGAGGCTTTGGCGGCTGCGCTGCGGATTTCCCTGCGAAGC 3749
Db 4042 ACCAACTACGCTAATTTGGTGGCGGCTTCGGTGGCGCTGCGCACTGGCAATTTGCTTACAAGC 4101
Qy 3750 GGGGGCATACCAACACCTTTACTCGAGCAGCGAGCAAAACCGGGCGGAGCGGCTATGT 3809
Db 4102 TGGGGGATCCCGCTCTTACTGCTTGAACAGTGATTAACCCGCGGCTGGGCTTATGT 4161
Qy 3810 GTTTGAGCAGAGTGGCTTTTACCTTCGATGCGGAGCCACCGGTGATCAACGATCCAGCGC 3869
Db 4162 CTACGAGGATCAGGGGTTTACCTTTGATGAGGCCCGACGGTTATCACCGATCCAGTGC 4221
Qy 3870 CATCGAAGATTGTTACGCTGGCAGGAAATCGCTCAGCGATTAACCTGCGAGCTGATGCC 3929
Db 4222 CATTAAGAACTGTTTGCACCTGGCAGGAAACAGTTAAAGAGATATGTGCAACTGCTGCC 4281
Qy 3930 GGTAAACGCCCTTCTATCGCCTGTCTCGGAAGATGGCAACAGCTTGAATACGACAATAA 3989
Db 4282 GGTACGCCGTTTACCGCCTGTGTTGGAGTCAAGGAGGCTCTTAAATACGATAACGA 4341
Qy 3990 TCAGCCGCTGCTGAGCAGCAGATCGCCACGTTCAATCCGCAAGATGTAGAAGGCTATCG 4049
Db 4342 TCAAAACCGGCTCGAAGCGCAGATTTCAGCAGTTTAAATCCCGCGATGTGGAAGGTTATCG 4401
Qy 4050 TCAATTTCTTCCTTATTCAGTGAAGTATTTAGAGAGGTTTATCTGAAACTCGGCACGGT 4109
Db 4402 TCAGTTTCTGACTATTTCACGCGCGGTTTAAAGAGGCTTATCTAAAGCTCGGTACTGT 4461
Qy 4110 GCGGTTCTGCGAGTGGCTGACATGCTCGCGCTGCGCGCAGTTGGGACGCTGCGAAGC 4169
Db 4462 CCTTTTATTCGTTTCAGAGACATGCTTCGCGCGCACCTCACTGCGGAACTGCGAGC 4521
Qy 4170 ATGCGCAGCGCTTCAGCAGATGTTGGGGAATTTTTCAGGACGATCATCTGCGTCAGGC 4229
Db 4522 ATGAGAGAGCGTTTACAGTAAGTTGCGAGTTACATCGAAGATGAACATCTGCGCCAGGC 4581
Qy 4230 GTTTTCTTCCTTCACATTCATTTGCTGGGCGGTAATCTTTTCAACGCTCATCTCATATAC 4289
Db 4582 GTTTTCTTCCTTCACATTCATTTGCTGGGCGGCAATCCCTTCGCGCACTCATCTCATATAC 4641
Qy 4290 CTTAATTCATGCTGAGCGTGAATGGGCGCTGTGTTTCCGCGCGCGGCAACGCGGC 4349
Db 4642 GTTGATACACGCGCTGGAGCGTGTGGGCGCTGTGTTTCCGCGTGGCGGCAACGCGGC 4701
Qy 4350 GCTGTGCGGCGATGGCGGACTGTTTCGAGGACTTGGGCGGCGGAGCTGTTTACTGAATGC 4409
Db 4702 ATTAGTTTCAGGGGATGATAAAGCTGTTTCAGGATCTGGGTGGCGAAGTCTGTTTAAACGC 4761
Qy 4410 CGAAGTGAAGCAGCTGGGAAACAGCGGCAATCGCAATAGCGCGTTCAGTTTAGAGGCGG 4469
Db 4762 CAGAGTGAAGCATATGGAACAGCAGGAAACAGATTGAAGCCGCTGCAATTAGAGACGG 4821
Qy 4470 ACGAGCTTCGATGCGCGCTGTGGCTTCAATGCGGAGCTGTTGCTGATACCTAGCAAA 4529
Db 4822 TCGCAGGTTCTGACGCAAGCGCTCGCTCAATGAGATGTGTTTCACTATCGGGA 4881
Qy 4530 ACTGTTTCGCAACCATTCGCTGGGCAATGAACGTCGCAATCGCTGAAGCGCTAAGCGCAT 4589
Db 4882 CCGTTTAAGCCAGCACCTTCGCGGTTAAGCGTCCAAACACTGCAAGCTAAGCGCAT 4941
Qy 4590 GAGCAACTCGCTGTTTGTACTTATTTTGGCTGATCAGCGCATGAAACAGCTCGCGCA 4649
Db 4942 GAGTAACCTCTGTTTGTGCTCTATTTTGGTTTGAATCACCATCATGATCAGCTCGCGCA 5001
Qy 4650 CCACACCGCTGTTTGGCCCGGCTTATCGTGAAGTGAATGAGATTTTCAACAGCAG 4709
Db 5002 TCACACGCTTGTTCGCGCGCGGTTACCGCGAGCTGATTCGAGGAAATTTTAAATCATGA 5061

QY 4710 CCAGCTGGCAGACGATTTTTCACCTTACTGCAACGCGCCCTGCAGCAGCGATCCGTCGCT 4769
Db 5062 TGGCCCTCGCAGAGGACTTCTCACTTTATCTGCAACGCGCCCTGTGTACGGAATTCGTCAC 5121
QY 4770 GGCACCGCCCGCTGCGCAGCTTTTATGTGTAGCGCGGTGCGCATCTCGGCACCGC 4829
Db 5122 GCGCCTGAAAGTTGCGCAGTTACTATGTGTGGCGCCGCTGCGCATTTAGGCACCGC 5181
QY 4830 TGACATCGACTGCAACAGAAAGGACCGCGTTGCGCGATCGAATTTTTCCTATCTGGA 4889
Db 5182 GAACTTCGACTGCACTGAGCGGTTGAGGGGCAAACTACGCGACCGTATTTTTCGCTACCTGA 5241
QY 4890 GCAGCACTACATCGCGGATTAAGTCAGCAATTAAGTGCACACAGATGTTTACCGCGTT 4949
Db 5242 GCAGCAATTAATCGCTTGGCTTACGGAGTCAGCTGGTCAACGACCGGATTTTACGCGCTT 5301
QY 4950 TGAATTTTCGCAACACCTGATGCCCATCACGCTCGCGGTTTTCGCTCGAGCCGATTTT 5009
Db 5302 TGAATTTTCGCAACACCTTAATGCCCTATCATGGCTCAGCTTTTCTGTGAGCGCGTTCT 5361
QY 5010 GAGCAAAAGCGCTGTTCCGCCGCAATAACCGCGATGCCGATATCAGCAATCTCTATCT 5069
Db 5362 TACCCAGAGCGCTGTTTCGGCGCATAAACCGCATAAACCATTAATCTCTACCT 5421
QY 5070 GTTGGTGGCGGTACGATCCAGCGCGCGCTGCGCGGTGATCGGTTGCGCCAGGC 5129
Db 5422 GGTGCGCGCAGGACACATCCCGCGCAGCGATTCCTGCGCTCATCGGTCGCGCAAGC 5481
QY 5130 CACCGCAGCGCTGATCTCGAGGATCGCGCGAATGAATCGACAGCTTTTACTTGAGCAA 5189
Db 5482 GACAGCAGTTTGATGCTGAGGATC---TGATTTGAATTAATCCGCTGTTACTCATCAT 5538
QY 5190 GTAAACCAACCATGCGCGTGGCTCGAAGATTTCCGCAACCGCCCAAGCTGTTTGAT 5249
Db 5539 CGCGTCAAAACGATGCGAGTTGGCTCGAAAGTTTTCGACAGCCTCAAAGTTATTTGAT 5598
QY 5250 GCACCAACCGCGCAGCAGCTGATGCTGTATGCTGTGTGCTGCTCACTCGCATGATGTG 5309
Db 5599 GCAAAACCGCGCAGCAGCTGATGCTCTAGCCTGCTGCGCGCATTTGTGACGATGTT 5658
QY 5310 ATTGATGGGCAACGCTGGCGAAGCGGCAGCAGCATGCGCTCGAAGCAGCGCAGGCA 5369
Db 5659 ATTGAGATCAGACGCTGGGCTTTAGGCCCGGCGAGCTTGCCTTACAAACGCGCGCAAA 5718
QY 5370 CGTATCAGCATCTGCAAAATTGAACCCCGCGCTTACAGCGCGCGCAGCATGATGAA 5429
Db 5719 CGTCTGATCAACTTGAGATGAACCGCGCAGGCTATGCGAGATCGCAGATGCAAGAA 5778
QY 5430 CCGCGGTTTAGCGGTTTTCAGGAGTGGCGATCAATTCACGAGTGGCGCAACATGGCG 5489
Db 5779 CCGCGGTTTTCGCGCTTTTCAGGAAGTGGCTATGGCTCATGATATCGCCCGGCTTACGCG 5838
QY 5490 TTTGATCATCTGAAGGCTTCGCTATGATGACGCAACGAAATTAACGCGAGCTTCGAT 5549
Db 5839 TTTGATCATCTGAAGGCTTCGCGATGATGATGACGCAAGCGAATACAGCCAACTGGAT 5898
QY 5550 GACACGCTCGTTACTGCTATCACTGCGCGCGTGGTGGTTTGTATGATGCGCGCGTA 5609
Db 5899 GATACGCTCGCTATTGCTATCAGTTGCGAGGTTTGTGCGCTTGATGATGCGCAATC 5958
QY 5610 ATGGGCGTGGCGACGAAGCGGTGCTGATCAGCCTCGGATTTAGGACTGGGTTCCAG 5669
Db 5959 ATGGGCGTGGGATTAACCCAGCTGGACCGCGCTGTGACCTTGGGCTGGCATTTTCAG 6018
QY 5670 CTCACATAACATTCGCGCGCAGATTTGAGAAGATGCGGAAATGCTGCTATCTGCGG 5729
Db 6019 TTGACCAATATTGCTCGCGATTTTGTGAGATGCGCATGCGGCGCGCTGTTATCTGCGG 6078
QY 5730 CAATCTGCTGATCAGCGGGATTAACGCGCGATACGCTGACTGACCGCAACATCGT 5789
Db 6079 GCAAGCTGGCTGAGCATGAAGTCTGAACAAAGAGATTAATGTCGCGCACCTGAAACCGT 6138
QY 5790 GCAGCGCTCGCTCACTGGCAGCGCGCTTTAGTGGCGGAGCGGAAACCTTATTATCACTCG 5849

RESULT 10

ACF67367_41/c

Continuation (42 of 57) of ACF67367 from base 4100001 (Photorhabdus luminescens nucleot

WP	Sequence	split	into	57	fragments	LOCUS	ACF67367	Accession	ACF67367
WP	Fragment	Name	Begin	End					
WP	ACF67367_00		1	110000					
WP	ACF67367_01		100001	210000					
WP	ACF67367_02		200001	310000					
WP	ACF67367_03		300001	410000					
WP	ACF67367_04		400001	510000					
WP	ACF67367_05		500001	610000					
WP	ACF67367_06		600001	710000					
WP	ACF67367_07		700001	810000					
WP	ACF67367_08		800001	910000					
WP	ACF67367_09		900001	1010000					
WP	ACF67367_10		1000001	1110000					
WP	ACF67367_11		1100001	1210000					
WP	ACF67367_12		1200001	1310000					
WP	ACF67367_13		1300001	1410000					
WP	ACF67367_14		1400001	1510000					
WP	ACF67367_15		1500001	1610000					
WP	ACF67367_16		1600001	1710000					
WP	ACF67367_17		1700001	1810000					
WP	ACF67367_18		1800001	1910000					
WP	ACF67367_19		1900001	2010000					
WP	ACF67367_20		2000001	2110000					
WP	ACF67367_21		2100001	2210000					
WP	ACF67367_22		2200001	2310000					
WP	ACF67367_23		2300001	2410000					
WP	ACF67367_24		2400001	2510000					
WP	ACF67367_25		2500001	2610000					
WP	ACF67367_26		2600001	2710000					
WP	ACF67367_27		2700001	2810000					
WP	ACF67367_28		2800001	2910000					
WP	ACF67367_29		2900001	3010000					
WP	ACF67367_30		3000001	3110000					
WP	ACF67367_31		3100001	3210000					
WP	ACF67367_32		3200001	3310000					
WP	ACF67367_33		3300001	3410000					
WP	ACF67367_34		3400001	3510000					
WP	ACF67367_35		3500001	3610000					
WP	ACF67367_36		3600001	3710000					
WP	ACF67367_37		3700001	3810000					
WP	ACF67367_38		3800001	3910000					
WP	ACF67367_39		3900001	4010000					
WP	ACF67367_40		4000001	4110000					
WP	ACF67367_41		4100001	4210000					
WP	ACF67367_42		4200001	4310000					
WP	ACF67367_43		4300001	4410000					
WP	ACF67367_44		4400001	4510000					
WP	ACF67367_45		4500001	4610000					
WP	ACF67367_46		4600001	4710000					
WP	ACF67367_47		4700001	4810000					

WP	ACF67367_48	4800001	4910000	
WP	ACF67367_49	4900001	5010000	
WP	ACF67367_50	5000001	5110000	
WP	ACF67367_51	5100001	5210000	
WP	ACF67367_52	5200001	5310000	
WP	ACF67367_53	5300001	5410000	
WP	ACF67367_54	5400001	5510000	
WP	ACF67367_55	5500001	5610000	
WP	ACF67367_56	5600001	5648894	
Query Match				19.6%; Score 1374.8; DB 10; Length 110000;
Best Local Similarity				62.0%; Pred. No. 0;
Matches 2211; Conservative				0; Mismatches 1347; Indels 8; Gaps 2;
Qy	2532	AATACGATGTGATTTTGGTTCGGTCTGGACTGGCGAATGGCTTGATTTGGCTGGCTGGCTGC	2591	
Db	23257	ACTGGGATCTGATTTCTGGTGGGAGGAGGATGGCCAAACGGACTGATGCCAATGGTTTCC	23198	
Qy	2592	GTCAATTGGACGACAACTGAAATGCTCTCTCGAGAGCGATGCGCATCCGCAGGCA	2651	
Db	23197	AGCAGTGTAACACACATCTGCGAGTGTGCTTATTGAAAACACAGAAACAATAGGCGCA	23138	
Qy	2652	ATCATACCTGGTTCGTTTCATCACAGCGATCTCAGGCCCGAACAACTTCGCTGGCTGCAAC	2711	
Db	23137	ATCACAGTGGTCAATTTCAATCAATGATCTTACTGAGCGGAAACATGATGSGATAGCAC	23078	
Qy	2712	CGCTGATTACCGTGCGTGTGTCAAGTTATCAGGTGCTTTTCTCGCTGCGCCGCAATC	2771	
Db	23077	CGCTGATTTACTATCGCTGTGTGAGTTAGACGTCAATTTTCCAGCATTTCAACGCACAT	23018	
Qy	2772	TGGACGGGGATTAATGTTCCATCGCATCAGGCGAATTTGCCCGCCATCTTTACGGCGGA	2831	
Db	23017	TGCCACATTCATATTTCAATCAGTATCACATCCAAACATTTTGCAAGCATCTCCATGCTAAT	22958	
Qy	2832	TGGGTGAGATCTGTGGGACAAACACAGCGGTACACAGGTAAACCCACGAGGTGACGC	2891	
Db	22957	TGGCGGAACGTATACAGACTCGTTTATGTTGTCAGGAGCTGACTCCACAGAAAGTTTACT	22898	
Qy	2892	TGGCGGATGGCGGTGAACTTGTCTGCGCAAGTGTGTGATGATGGTTCGCGCTTCGACCGGA	2951	
Db	22897	TACAGGACGGCTCGTCTCTAAGTGCTGGCGCAGTCAATCATGGGAGAGGCTGGCGACCA	22838	
Qy	2952	CGCCACATCTGACGTGGGTATCAGGTGTTTCTTGGAACAGATGGCAGCTGGCGCAGC	3011	
Db	22837	GACCAITTTATGGGAGTGGCACCCAGGCAATTTTGTGTCAGAAATGGGAGCTGGAAAGT	22778	
Qy	3012	CGCAGCGCTGCACAGCCGATCTGATGATGCCACCGTCTGATCAGCAATGCAAGCGGTTATC	3071	
Db	22777	CGCACTCTTTAAACCCACCCGATTTTATGATACTAGTGTGGACAGGATACAGGTATC	22718	
Qy	3072	GTTTTGTCTACCGCTCCGCTCAGCGCCGATCGGCTATTGATTGAAGTACCCATTAGC	3131	
Db	22717	GATTTATCTATGTCTCCGTTCTCTCAACTCGTCTGCTGTATAGAGGACATCTCATTCG	22658	
Qy	3132	TTAACCCAGCCGCTGGCGGGAACACCCGCTCGTCAGCACATCGCGCACTATGCAATC	3191	
Db	22657	TTCATCGGGGCGCACTGATAAGGCTTTGTTCGACGGCTACTATTCGACAGTACGGGAGA	22598	
Qy	3192	AGCAAGCTGGACGCTGAGTACGCTGTGCTGGAAGACGACCGGCATATTACCGATTACCC	3251	
Db	22597	AACATGATGAAACTGGTAAACTCATTCGAGAAGAGCGGTTGCTCTTCCAAATTACAC	22538	
Qy	3252	TGAGCGGCAACATCGATCGATTCTGGCAACAGACGCGCGCCAAAGCTGACGGCCCTGC	3311	
Db	22537	TTACGGGAGATTTTACCTCTTTCTGGGCAAGCTTAGCAGACGCCCACTCTGGGTTAC	22478	
Qy	3312	GGCGCGGCTGTTTTCATGCCACCAACCGGTACTCTTTCGCTCGCGCTGGGCTAGCGG	3371	
Db	22477	GTGCACTTTGTTTTCACCCCAACAGGCTACTCTCTGCCACACGCTATTTCGTTGGCAG	22418	
Qy	3372	AGTTGGTACGACGCTGTGGCCACCGATGCCCTCAGCTCAGCGCAACATATCGAACGCT	3431	
Db	22417	ATCGTATTGTTGCTCTGCGGAGCTTACCGATACCTCTTATTATTATTAACCTCAGGAT	22358	

QY 4507 GAGTGGTGCATACCTACGACAAACTGCTTCGCCACCATCCGCTGGCAATGAACGTGGC 4566
DB 21277 GATGTTATTTATACATTAAGACACCTGTTAGTGACATCCGCTCGGAATAGCCCGCT 21218
QY 4567 ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTGTACTCTATTTTGGCTGTAAT 4626
DB 21217 CGAACACTGGTGGTAAGCAATGAGCAACTCACTGTTGCTGCTTTATTTTCGGTCTGAAC 21158
QY 4627 CAGCGCATGAACAGCTCGCGCACCAACACCGTCTGTTTGGCCCGCTTATCGTAGTTG 4686
DB 21157 CATCACCATACGCAATTAGCCCATCATACAGTTTGTCTGGTCCAGCTATAAGGAATC 21098
QY 4687 ATCGATGAGATTTTCAACAGCAGCAGCTGCGCAGACGATTTTCACTTTACCTGACCGG 4746
DB 21097 ATCGAAGATATTTTATCATGACCGACTGTGAGAGGATTTTCACTCTATCTCATGCC 21038
QY 4747 CCTGCGCAGCAGGATCCGCTCGCTGGCACCGCCCGCTGGCAGCTTTTATGTTTAGCG 4806
DB 21037 CCTCTGTTACTGACCCATCCCTTGTCTCCCGAGGATGCGAAGTTACTAGCTTCTGGCC 20978
QY 4807 CCGGTGCGCATCTCGGCAACCGCTGACATCGATGCAACAGGAAGGACCGCGTTGCGC 4866
DB 20977 CCGTGTACCAATTTGGGACAGCTAATTTAACTGGGATATAGAGGGGCGACGCTTACGT 20918
QY 4867 GATCGAATTTTGTCTATCTGAGCAGCACTACATGCGCGGATTAGCTCAGCAATTAGTG 4926
DB 20917 GACCGTATTTTCGCGTATCTGAAAGTACTATATGCCCGTCTGCTCAAGCAGTTAGTC 20858
QY 4927 ACACACAGAAATGTTTACGCGCTTTGATTTTCCGCACACGCTGCATGCCCATCACGCTCG 4986
DB 20857 GTTCATCGTATTTTACCCATTTGATTTTGTGACCACTCAATGCCCATCTCGGTTG 20798
QY 4987 GCGTTTTCGCTGAGCGGATTTGACGAAAGCGCTGTTCCGCCCGCATAACCGCGAT 5046
DB 20797 GCCTTCTCTTTGAACCGTTGCTGACACAAAGCGTTGTTCCGACCACACAATCGGAC 20738
QY 5047 GCGGATATCAGCAATCTCTATCTGTTGGTGGCTAGCGATCCAGCGCGGCGTGGCC 5106
DB 20737 AACCGCATTTGATATCTGATCTCGTGGCGCAGGACGACATCTCTGGAGCAGGTATCCC 20678
QY 5107 GCGGTGATCGGTTCCGCCAAGCGCACCGCAGCGTGATCTCGAGGATCGCGCCGAATGA 5166
DB 20677 GCGGTGATTTGATCAGCAAAAGCCACCGCCCACTTGATGTTTAGAGGATATCGCTAATGA 20618
QY 5167 ATGCAAGCGCTTTACTTTGAGCAAGTAACGCAACCATGCGGTGGCTCGAAGATTTG 5226
DB 20617 ATCTGCGC-...TACTTAAACAGTTTACTCAGATAATGAGCAAGGCTCAAAAAGTTTG 20561
QY 5227 CCACCGCGCAGCTGTTGATGCACGACGCGCGCAGCAGCTGATGCTGTATGCT 5286
DB 20560 CCAGTGTCAACCGACTTTTCGATACAGCAACACGACACAAATGATGCTGTACGCT 20501
QY 5287 GGTGTGCTCACTCGCATGATGATTGATGGGCAACGCTGGCGAAGCGCGCACGACG 5346
DB 20500 GGTGGCGTTATTTGATGATATATAGAGCGGCAAGAGCTGGGAGACAAATAAGCAGTG 20441
QY 5347 ATGCGGTGCAAGACGCGCAGGACGATPACGACATCTGCAATTTGAAACCCCGCGCT 5406
DB 20440 TTGATAAGTATAGCGCGCCGAGAACTTCAAAATGCTCAATACCTGCAAAAGCAGGCTT 20381
QY 5407 ACAGCGCGCGCACATGAGTAACCGCGTTTAGGGCGTTTACGAGAGTGGCGATCATTC 5466
DB 20380 ACAGCGGTTTGGCGATGACCGCAACCTGCTTTTGCAGCCTTTTCAGACAGTAGCTGTAGTA 20321
QY 5467 ACCAGCTGCGCAACAACTGGCGTTTGTATCATCTGAAGGCTTCGCTATGATGACACGA 5526
DB 20320 ATGAAATCCTCAGCAACAGGCTTTCGAACATCTGGAAGATTCGCAATGGATGACTCT 20261
QY 5527 ACGAACATTTACGAGCTTTGATGACACGCTCGGTTACTGCTATCACTGCGCGGCGTGG 5586
DB 20260 GTGAGCGCTTACAGAACGTTTGGATGATCTCTGAGATATTGCTTACCATGTCCGGGAGTAG 20201
QY 5587 TCGGTTTGTATGATGGCGCGCTAATGCGGCGTGGCGACGAAGGGTGTCTCGATCAGGCT 5646

DB 20200 TTGGGCTGATGATGCAAGGGTTATGGGGGTTTCGAGCGGAGTGTGCTGGATCGGGCT 20141
QY 5647 GCGATTTAGGACTGGCGTTCAGCTCACTAAATTCGCGCGGACATTTGTAGAAATGCGG 5706
DB 20140 GCGATCTAGGAATGCTTTTCACTAACTAATTCCTCGGATATTTAGAGGATGCA 20081
QY 5707 AAATGGTGGCTGCTATCTGCGCAATCTGCGCTCGATCAGCGCGGATTAACGCGCGATA 5766
DB 20080 AGGCGGGCGTGTCTATCTTCGCTAGAAATGGCTTTCACCAAGAGGGCTAATGCCAGACA 20021
QY 5767 CGCTGACTGCACCGCAACATCTGCGAGCGCTCGCTCACTGCGACGCGGTTTGTAGTGGCG 5826
DB 20020 CACTTATTTATACCGAAACCGTCTCGCTAGCTCGGGTGGCATCCCGATGATTGTGG 19961
QY 5827 AGCGGAAACCTTATTTATCACTCGGCGCATCCGCTTTTACCGGTTTACCGCTCGGCTCGG 5886
DB 19960 AAGCGGATCTTATTTATACCTCGCTCGACAGGGCTTGTGTTTACCTTTGCGTTTCA 19901
QY 5887 CGTGGCCCATCGCTACGCTCGCGCTTTTATCGGAAATTTGGCTCAAAAGTTCAAGCAG 5946
DB 19900 CCGGTGCTATCGCTTCTGCTCATGTATCTATCTGTAATCGCATTAAGTGCAACAG 19841
QY 5947 CGGTGTGCGCGCTGGGATTCACGCGAGCGCACCAAGTAAAGTGAAGAACTGGCGCTGC 6006
DB 19840 CAGGAGTACGGGCTTGGGAATGACAGAAAGAACTAATCGTGGAGAGAGTTTACCTTAT 19781
QY 6007 TGGTGAAGGGGCGAGTTTGGCGATCACTTCGCTGTGTCTGCTCTTGAACCGCTCGG 6066
DB 19780 TAATGGCTGGCGCAATGAAGCGCTTATATCTCGGTGGCGATAACAAGCCTCGGATC 19721
QY 6067 CTGCTGTGCGAGCGCTCTCTGTTGA 6092
DB 19720 CCAATTTGCGCAGCGACCTCATTA 19695

RESULT 11

ACF65388_06
Continuation (7 of 13) of ACF65388 from base 600001 (Photorhabdus luminescens nucleoid
WP Sequence split into 13 fragments LOCUS ACF65388 Accession ACF65388

WP	Fragment Name	Begin	End
WP	ACF65388_00	1	110000
WP	ACF65388_01	100001	210000
WP	ACF65388_02	200001	310000
WP	ACF65388_03	300001	410000
WP	ACF65388_04	400001	510000
WP	ACF65388_05	500001	610000
WP	ACF65388_06	600001	710000
WP	ACF65388_07	700001	810000
WP	ACF65388_08	800001	910000
WP	ACF65388_09	900001	1010000
WP	ACF65388_10	1000001	1110000
WP	ACF65388_11	1100001	1210000
WP	ACF65388_12	1200001	1225559

Query Match 19.6%; Score 1374.8; DB 10; Length 110000;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 2211; Conservative 0; Mismatches 1347; Indels 8; Gaps 2;

QY 2532 AATACGATGCTGATTTTGGTTCGCTGCTGAGCTGGCAATGGCTTGTATGCGCTGCGCTCGC 2591
DB 101680 ACTGGGATCTGATTTCTGTTGGAGGAGGATTTGGCCACCGGACTGATCGCAATGCGTTTCC 101739
QY 2592 GTCAATTGCGACCAACCTGAAATCCCTGTTGCTGGAGAGCGATGCGCATCCGCGAGGCA 2651
DB 101740 AGCAGTGAACCAACATCTGCGAGGTGTTGTTATTTGAAACACAGAAACAATAGAGGCA 101799
QY 2652 ATCACTGCTGCTGTTTATCAGCGGATCTCAGCGCGGCAACAATTCGCTGCTGCTGCAAC 2711
DB 101800 ATCACACGCTGCTGCTTCTTCACTCAACATGATCTTACTGAGCGGGAACATGATGGATGAC 101859
QY 2712 CGCTGATTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2771

Db 101860 CGCTGATTACCTATCGCTGGTACAGGTTACAGGTCATTTTTCAGCAATTTCAACGCACAT 101919
Qy 2772 TGGACGGGATTAATGTTCCATCGCATCAGGCGAATTTTTCGCCGCAATCTTTACCGCGCA 2831
Db 101920 TGCCACATTCATTTTCAGTATCAATCCCAACACTTTTGCAAGCATACTCCATGCATATT 101979
Qy 2832 TGGGTGACGATCTGTGGCAAAACAACAGCGGTACAACAGGTAAACCCACGAGGTGAGCG 2891
Db 101980 TGGGCGAACGTATACAGACTCGTTTATTTGGTACAGGAGTCACTCCACAGAAAGTTTACT 102039
Qy 2892 TGGCGGATGGCGTGAATCTGTCGCAAGTGGTGAATGATGCTCGCGGCTGCGAGCGCA 2951
Db 102040 TACAGACCGCTCGCTCTAAAGTGTGGCGAGTCAITGATGGAGAGGCTGGCGACCAG 102099
Qy 2952 CGCCACATCTGCAGCTGGGTTATCAGGTGTTTCTTGGACAAGATGGGAGCTGGGCGAGC 3011
Db 102100 GACCAITTAAGGAGTGCCACCCAGGCAATTTTGGTCAGGAATGGGAGCTGGAAAGT 102159
Qy 3012 CGACCGCTGACGAGCGATCCTGATGATGCCACCGTCGATCAGCAAGCGGTTATC 3071
Db 102160 CGCACTCTTTAAACCCACCCGATTTTAAATGGATAGTGTGGGACAGGATACAGTTATC 102219
Qy 3072 GTTTGCTACAGCTGCGGCTCAGCGCGATCGGCTATTTGATGAAGATACCCATTAG 3131
Db 102220 GATTTATCTATGCTGCGGCTCTCTCACTCGTCTGATAGAGACACTCAATTAG 102279
Qy 3132 TTAAACGACCGCGCTGGCGGAGAACACCGCTCGTCAGCACATCGCGACTATGCCAATC 3191
Db 102280 TTGATCGGGGCGCACCTGATAGGCTTTGTGCGAGGCTACTATCGCAGATGACGGAGA 102339
Qy 3192 AGCAAGCTGGAGCGCTGAGTACGCTGCTGGTGAAGAGCACCGCATATTACCGATTACCC 3251
Db 102340 AACATGATGGAACCTGGGTAACTCATTCGAGAGAGCGGTTGTCTTCCAAATTACAC 102399
Qy 3252 TGAGCGCAACATCGATCGATTCTGGCAACAGCAGCGCGGCGAAGGCTGACGGCGCTGC 3311
Db 102400 TTAGCGGAGATTTTACCTCTTTCTGGGCAACAGCTAGCAGGACAGCCACCTGTGGGTAC 102459
Qy 3312 GCGCGGCTGTTTCATGCCACCAACCGTTACTCTTGGCTGCGCGTGGGCTAGCGG 3371
Db 102460 GTGAGCTTTGTTTACCCCAACAGGCTACTCTGCGCACAGCTATTCGGTTGGCAG 102519
Qy 3372 AGTTGTAGCAGCGCTTTGCCACCGATGCCCTCAGCTCAGCCAAACATATCGAACGCT 3431
Db 102520 ATCGTATTTGCTCTCGCGAGCTTACCGATACCTCTTATTCATTAACCTCAGGGATT 102579
Qy 3432 TTGCGGCTGACAGTGGCGGCAACAGGATTTTTCGCTGCTCTAAACCGCATCTGTTT 3491
Db 102580 ACGCACGACAAACAGTGGCAACACAGCGCTTTTTCGCTCTTAAATCGCATGCTCTTC 102639
Qy 3492 TGGCGGTTAAGCCGACAGCGCTGGCGGCTGATGCAACGTTTTTACCGGCTCGATGCCG 3551
Db 102640 TCGTGGGATCCACAAACAGTTGGCAGGTATGCAACGTTTCTATCAACTTCCCAA 102699
Qy 3552 GGTAAATPAGCGCTTTTACCGCGGCAACGTCGCGCTGCGGATAAACCGGATTTCTGT 3611
Db 102700 ATCTGATTCACGTTTTTATCGGAGCAACTTAATTCGTCGACAGGCGCGATTCTCA 102759
Qy 3612 GCGGCAAGCGCGGCTGCTGATGAGCGCTGCGCGGCTTTGAA-----TTCTGT 3666
Db 102760 TAGGTAACACACAGTGGCGGATTAAGGTTGCTCTAAAGGCNAATGTTTAAACAACAAGA 102819
Qy 3667 CGAACAGGGAAGAAAATGAACGCACTTATGATGCGGAGGCTTTGGCGGCTG 3726
Db 102820 AGCTTCAGGTTTTTATTAATGATTAAGCGCTGGTAAATGGTGTGTTTTTGGTGGCTG 102879
Qy 3727 GGGCTGGGATTCGCTCGAGCGCGGATACCAACCTTTACTCGAGCAGCGAC 3786
Db 102880 GCACCTGGCAATAAGGCTCGAGTCTCGGGGATTCGACATGATTTTGGAGCAACGGAT 102939
Qy 3787 AAACCGGGCGAGCGGCTATGTTGAGGACAGTGGCTTTACCTTCGATGCGGACCC 3846
Db 102940 AAACAGGTGGACGCGCTTATGTTTATAAGGAACAGGGAATTCACCTTGTATGCGGCCCC 102999

Qy 3847 ACGTGATCACCGATCCCGAGCCCATCGAAGAGTTGTTTACGCTGCGAGGAAAAATCGCTC 3906
Db 103000 ACGTAAATCACCGCTCTTAATGTCTATTGAGAAATGTTTACCAGGCTGGTAAACGTATG 103059
Qy 3907 AGCGATTACGTCGAGCTGATGCGGTAAAGCCCTTCTATCGCTGTGCTGGGAAGATGCG 3966
Db 103060 GCTGATTACGTCGATTTACTTCCCGTACACCCCTTTTATCGGCTCTGTTGGGAGTCGGGT 103119
Qy 3967 AAACAGCTTGAATPACGACAAATATCAGCGCTGCTGGAGCAGCATGCCACGTTTCAAT 4026
Db 103120 AAGATGTTGATPACGACAAATGATCAGCAACATCTGGAAGCACAATCCATACGTTCAAT 103179
Qy 4027 CCACAAGATGTAGAAGGCTATCGTCAATTTCTTGCGCTATTCACTGTAAGATTTTAGAGAG 4086
Db 103180 CCAAGAGATGTTAAACGGTATCGACGCTTTCTAGACTATTCCCGGAAGCCCTTTAATGAA 103239
Qy 4087 GGTATCTGAAATCGCGACGGTGGCGTTTCTGAGGTGCGTGACATGCTGCGCGTGGCG 4146
Db 103240 GGCTACCTGAACTTGGAAACGGTGGCGTCTCTCTTTTCGGGACATGTTGAGTGGGCT 103299
Qy 4147 CCGCAGTTGGAGCTCTGCAAGCATGGCGCAGCGCTCTACAGCATGGTGGCGGAAATTTATT 4206
Db 103300 CCACAGTTGATAGCTCTACAGCATGGCGTAGTGTATATAGTCAGGTGGCAGCGCTTTATC 103359
Qy 4207 CAGCAAGCATCTGCGTCAAGCGTTTTCTTCCACTCATTTGCTGGTGGCGGTATCTCT 4266
Db 103360 AAGGATGAGAGCTTGGCCAGGCAATTTCTATTTCACTCGCTATTTAGTAGGAGTAACCC 103419
Qy 4267 TTTTCAACGTCATGATCTATCTTAATTTATCATGCTGAGCGTGAATGGGCGGTGTGG 4326
Db 103420 TTTGCTGATCTCTATCTATATCTTTGATCCAGCATGGGAGTGG 103479
Qy 4327 TTTCCGCGCGCGCACCGCGCTGGTGCAGGGCATGGCGGACTGTTTCGAGGACTTG 4386
Db 103480 TTTCCGCGAGGAGAACCGCGCTCTGTTGAGGCAATGGTGAATTTGTTTACTGACANT 103539
Qy 4387 GCGCGGAGCTGTTACTGAAATGCGAAGTGAAGCGAGCTGGAACACGAGCGCAATTCGCAAT 4446
Db 103540 GGTGAGAAATTTGAGCTTTAATGCAAGGTAAACGCTTCCACCCATGGCAATCGGTC 103599
Qy 4447 AGCGCGTTTACGTTAGAGGCGGACGAGCTTCGATCGCGCTGTGGCTTCAATGCC 4506
Db 103600 ACAGGTGTCATTTAGCTGATGGAGCAATATGATGATGAGCTGTGGCTCAATGCC 103659
Qy 4507 GACGTGTGCTACTTACGACAACTGTTCCGACCATCGCTGGCAATGAAACGTCG 4566
Db 103660 GATGTTATTCATCTTAAAGCCTGTAGTCAGCATCCGTCGGAATAGCCCGCT 103719
Qy 4567 ACATCGCTGAAGCTAAGCGCATGAGCAACTCGCTGTTGTTGTTGTTGCTTATTTTGGCTGAAT 4626
Db 103720 CGAACACTGGTGGTAAACGATGAGCACTCACTGTTGCTGTTTATTTTCGGTCTGAAC 103779
Qy 4627 CAGCGCATGAACAGCTCGCGACCAACCGCTGTTTGGCCCGCTTATCGTAGTTG 4686
Db 103780 CATCACATACGCAATTAAGCCCATCATACAGTTTGTCTTGGTCCAGCTTAAGGAATC 103839
Qy 4687 ATCGATGAGATTTTCAACAGCGCGAGCTGGCAGAGATTTTTCATCTTACTTCGACGCG 4746
Db 103840 ATCGAAGATATTTTTTATCATGACCGACTGTGAGGATTTTTCATCTATCTCCATGCC 103899
Qy 4747 CCCTGACGAGCATCCGTCGCTGGCACCGCGCTGGGAGCTTTTATGTTGTAGCG 4806
Db 103900 CCCTGTTTACTGACCCATCCCTTCTCCCGAGGATGCGAAGTTACTGTTCTTGCC 103959
Qy 4807 CCGGTGCGCATCTCGGCACCGCTGACATCGACTGCAACAGGAAGGACCGGCTTGGCG 4866
Db 103960 CCGTGTACCAATTTGGGACAGCTAATTTAAACTGGATATAGAGGGCCAGCTTACGT 104019
Qy 4867 GATCGAATTTTGTCTTATCTGGAGCAGCACTACATGCCGGGATTAAGTCAAGAAATAGTG 4926
Db 104020 GACCGTATTTTTCGCTATCTGGAAAGTACTATATGCCCGCTCTGCTCAAGCAGTTAGTC 104079

Db 2280 TTGATCGGGGCCACCTGATAGGCTTTTTCGAGGCTACTATCGCAGATACGGAAGA 2339
QY 3192 ACBAAGGCTGACCGCTGAGTAGCTGCTGCTGTAAGACACGGCATATTACGATTACCC 3251
Db 2340 AACATGGATGGAACATGGGTAACTCATTCGAGAAGAGAGCGGTGTCTTCCAAATTAC 2399
QY 3252 TGACGGGCAACATCGATCTCTGCAACACGACGCGCGGCAAGCTGACAGCGCCCTGC 3311
Db 2400 TTACGGGAGATTTTACTCTTCTGCGCACAGCTAGCAGGACAGCCACCTGTGGTTAC 2459
QY 3312 GCGCGGGCTGTTTATCATGCCACACCGGTTACTCTGCGGTCGCGCTGCGGCTAGCGG 3371
Db 2460 GTGCAGCTTTGTTTCAACCCCAACACAGGCTACTCTCTGCCACACAGCTATTGCGTTGGCAG 2519
QY 3372 AGTTGGTAGCAGCGCTGTGCCCCACCGATGCCCTCAAGCTCAGCCAAACATATCGAACGCT 3431
Db 2520 ATCTGATTTGCTCTGCGCGAGCTTACCGATACCTCTCTTATTCATTACCTCAGGAT 2579
QY 3432 TTGCCCCGTGACAGTGGCGCAACAGCGATTTTTCGCTGCTAAACCGCATGCTGTGTTT 3491
Db 2580 AGCAGACAAACAGTGGCAACACAGCGCTTTTTCGCTCTTAAATCGCATGCTCTTCC 2639
QY 3492 TGGCGGTTAAGCGCAGCAGCGCTGGCGGTGATGCAAGTTTATACCGGCTCGATGCGG 3551
Db 2640 TCGCTGGGATCCACAACACGTTGGCAGGTAATGCAAGCTTTCTATCAACTTTCCCAA 2699
QY 3552 GGTAAATTAGCCGCTTTTACGCGGGCAACTGCGCTGCGCATAAACCGCATTTCTGT 3611
Db 2700 ATCTGATGCAAGTTTATGCGGAGCAACTTAAATCCGTGCAAGGCCGGAATCTCA 2759
QY 3612 GCGGCAAGCGCGGCTGCCATCGGTGAAGCGCTGGCGGCTGTGAA-----TTCGT 3666
Db 2760 TAGGTAACACCAAGTGGCGGTAAGGGTCTCTAAAGGCAATGTTAAACAAACAAGA 2819
QY 3667 CGAACAGGAAGAAAATGAACCGCACTTATGATGCGCGCAGGCTTTGCGGCGCTG 3726
Db 2820 AGCTTCAGGCTTTTATTAATGATTAAGCGCTGGTAATTTGGTGTGCTGTTTGGTGGCTG 2879
QY 3727 GCGCTGGCATTCGCTGCAAGCGGCGGCATACCAACCACTTACTCGAGCAGCGGCAC 3786
Db 2880 GCATGGCAATAGGCTCAGTCTGCGGGATTCGACATGATTTTGGAGCAACGGAT 2939
QY 3787 AAACCGGGGAGCGCGCTATGTTTGAAGCAGGTGGCTTTTACCTTCGATCGCGACCC 3846
Db 2940 AAACAGGTGGAGCGCTTATGTTTATAAGGAACAGGGATTCACCTTTGATCGCGGCC 2999
QY 3847 ACGGTGATCACCGATCCAGCGCATCGAAGAGTCTTCACGCTGGCAGGAAATCGCTC 3906
Db 3000 ACCGTAATCACCGCTCTTAATGTCATGAAGAAATGTTTACCAGGCTGGTAAACGTATG 3059
QY 3907 AGCGATTACGTCAGCTGATGCGGTAACGCCCTTCTATCGCTGTGCTGGGAAGATGCG 3966
Db 3060 GCTGATTACGTCGATTTACTTCCGTTACACCTTTTATCGGCTCTGTTGGAGTCGGGT 3119
QY 3967 AAACAGCTTGATTACGACATATACCGCGCTCTGGAGCAGCAGATCGCCAGTTCAAT 4026
Db 3120 AAGATGTTTGAATACGCAATGATCAGCAACATCTGGAAGCAAAATCCATAGTTCAAT 3179
QY 4027 CCGCAAGATGTAAGGCTATCTGCAATTTCTTTCGCTATTTCAGCTGAAGTATTAGAGAG 4086
Db 3180 CCAAGAGATGTAACGGGTATCGACGCTTCTAGACTATTCGCGGAGCCTTAATGAA 3239
QY 4087 GGTATCTGAACTCGGCAACGCTGCGGTTTCTGAGGTGGCTGACATGCTGCGGTCGGG 4146
Db 3240 GGCTACCTGAACTTGGAAACGCTTCTCTCTCTTTTCGCGACATGTTGAGTGGCGCT 3299
QY 4147 CCGCAGTTGGAGCTGTCGAAGATGCGCAGCGCTCTACAGCATGTGTGGCGAATTTAT 4206
Db 3300 CCAAGTTGATAGCTCTACAGCATGGCGTAGTGTATAGTCAGTTGGCAGCCTTTATC 3359
QY 4207 CAGGACGATCATCTGCGCTCAGGCGTTTCTCTCCACTCATTTGCTGTGGCGGTAATCT 4266
Db 3360 AAGGATGAGAGCCTGGCCAGGCATTTTCACTTCGCTATTGATAGGAGTAACCC 3419

QY 4267 TTTGCAACGTCATCGATCTATACCTTAATTCATGCTGAGCGTGAATGGCGGCTGTGG 4326
Db 3420 TTTGCTGCACTTCTATCTATATACTTTGATCCACGCACTGGAACGAGATGGGAGTCTGG 3479
QY 4327 TTTCCGCGCGCGGCGACCGCGCGCTGTGTCAGGGCATGCGCGACTGTTTCGAGGACTTG 4386
Db 3480 TTTCCGCGAGGAGAACCAAGCGCTCTGGTTGAGGCAATGGTGAATTTGTTTACTGACATT 3539
QY 4387 GCGGCGAGCTGTTTACTGAAATGCGAAGTGCAGCTGAGAAACACGAGCGCAATCGCAT 4446
Db 3540 GGTGAGAGAAATGAGCTTAAATGCAAGTAAACGCTTACACACCATGCAATCGGGTC 3599
QY 4447 AGCGCGCTTCAGTTAGAGGCGGACGACGCTTCGATCGCCGCTGTGGCTTCCAAATGCC 4506
Db 3600 ACAGTGTCCAAATTAGCTGATGACGAAATATGACATGTGACGTTGTGGCTCAAATGCC 3659
QY 4507 GACGTGTGATACCTTACGACAACTGCTTTCGCCACCATCGCTGGCAATGAACGTCG 4566
Db 3660 GATGTTATTATACCTATAAGCACCTGTAGGTGAGCATCCGTCGGAATAGCCCGCT 3719
QY 4567 ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTGTACTCTATTTTGGCCTGAAT 4626
Db 3720 CGAACACTGGTGTGTAACGAATGAGCAACTCACTGTTGCTGTTTATTTTCGGTCTGAAC 3779
QY 4627 CAGCGCATGAACAGCTCGCGCACCAACCGCTCTGTTTTCGCCGCGCTTATCGTGATTG 4686
Db 3780 CATCACCATACGAAATAGCCCATCATACAGTTTGTCTTCCACGCTAAGGAACATC 3839
QY 4687 ATCCATGAGATTTTCAACAGCAGCGACGTCGACAGATTTTTCACTTTACCTGCAACGCG 4746
Db 3840 ATCGAAGATATTTTATATCATGACGACTGTGAGAGATTTTTCACCTATCTCATGCC 3899
QY 4747 CCCTGACGACGCTATCGCTCGTGGCACCGCGCTGCGGACGCTTTTATGTTAGCTAGCG 4806
Db 3900 CCCTCTGTTACTGACCCATCCCTTCTCCCGAGGATGCGCAAGTTACTAGTTTCTGGCC 3959
QY 4807 CCGGTGCGGATCTCGGCAACGCTGACATCGACTGGCAACAGGAAGACCGCGCTTCGCG 4866
Db 3960 CCTGTACCATTTTGGGAGACAGCTAAATTTAACTGGGATATAGAGGGGCCACGTTTACGT 4019
QY 4867 GATCGAATTTTGTCTTATCTGAGCAGCAGCTACATGCGGGGATTCGTCAGCAATTAGTG 4926
Db 4020 GACGATATTTTCGGTATCTGGAAGAGTACTATATGCCCCGCTCTGCTCAAGCAGTTAGTC 4079
QY 4927 ACACAGAAATGTTTACGCGCTTTGATTTTCGACAGCTGCTGATGCCCATCAAGGCTCG 4986
Db 4080 GTTCATCGTATTTTACCCCAATTTGATTTTCTGTGACCAACTCAATGCCCCATCTCGGTTG 4139
QY 4987 GCGTTTTCGCTGGAGCGGATTTTTCAGCAAGCGCTGTTTCCGCCCGCATAAACCGCAT 5046
Db 4140 GCTTCTCTCTTTGAACCGTTGCTGACACAAAGCGTTGTTTCCGACCAACAATCGGGAC 4199
QY 5047 GCGGATATCAGCAATCTCTATCTGTTGGTCCCGGTACGCAATCCAGGCGCGGGGTCGCC 5106
Db 4200 AACCGCATTTGATAATCTGTAICTGTCGCGCGCAGCAGCATCTCTGGAGCAGGTAATCCC 4259
QY 5107 GCGGTGATCGGTTGCGCCAAAGCGCACCGCAGGCTGATGCTGGAGGATCGCGCGGAAGA 5166
Db 4260 GCGGTGATGAGTACAGAAAACCCCACTGATGTTGATGTTAGAGGATATCGCTAAATGA 4319
QY 5167 ATCGACAGCTTTTACTTTCAGCAAGTAACGCAAAACCATGCGGCTGGGCTCGAAGAGTTTCG 5226
Db 4320 ATCTGCGG---TACTTAAACAGTTACTCAGATAATGAGCAAGGCTCAAAAAGTTTTG 4376
QY 5227 CCACCGCGCCCAAGCTGTTTGAATGACCGAGCGCGCGCAGCAGCTGATGTTATGCTG 5286
Db 4377 CCAAGTGTCAACCGACTTTTTCGATACAGCAACACGACACAGCAATGATGTTGTCGCT 4436
QY 5287 GGTGCTGCTCACTGCGCATGATGTTGATGGCAAAACGCTGGCGGAAGCGCGCACGACG 5346
Db 4437 GGTGCGGTTATTTGATGATATATATAGACGGGCAAGAGCTGGGAGACAAATAAGCAGTG 4496

QY 5347 ATGCGTGAAGACGGCAGGACGATGATGACGATCTGCAATTTGAAACCCGCGCGCT 5406
 Db 4497 TTGATAAGTATAGCCCGCGAGAACTTCAATGCTGCAATACCTGCAAGACGAGGTT 4556
 QY 5407 ACAGCGGCGGCACATGATGAACCGGCTTTAGGCGGTTTCAGGAAGTGGCGATCAATC 5466
 Db 4557 ACGACGGTTTGGCGATGACCGAACCCTGTTTGGACGCTTCAGACAGTGTCTGAGTA 4616
 QY 5467 ACCAGCTGCGCAACAACTGGCGTTTGTATCTGGAAGGCTTCCTATGAGATGACGCA 5526
 Db 4617 ATGAATCCCTCAGCAACAGGCTTCGAACATCTGGAAGGATTCGAATGATGATCT 4676
 QY 5527 ACGACATTACGCGAGCTTCGATGACAGCTGCTGCTATCATCTGCGGCGGCG 5586
 Db 4677 GTGAGCTTTACAGACGTTGGATGATCTCTGAGATATGCTTACCATGTCGCGGAGTAG 4736
 QY 5587 TCGGTTTGTATGATGCGCGGCTTAATGCGGCTGCGGACGAGCGGTGCTCATACAGCT 5646
 Db 4737 TTGGGCTGATGATGCGAAGGTTATGGGGTTTCGAGGCGAGTGTCTGATCGGGCT 4796
 QY 5647 GCGATTTAGGACTGGCGTTTCAGCTCACTAACTTGGCGGACATTTGAGAGATGCGG 5706
 Db 4797 GCGATCTAGGAATGCTTTTCACTTAACTTAACATGCTCGGATATTTAGAGATGCGA 4856
 QY 5707 AAAATGCTGCTGCTATCTGCGGCAATCTGCTGATCAGCGGATTAACGCGGATA 5766
 Db 4857 AGCGCGGCGTGTCTATCTTCGCTAGAAATGGCTTTCACCAAGAGGCTAATGCCAGACA 4916
 QY 5767 CCGTCACTGCAACGCAACATCTGCGAGCGCTCGCTCACTGCGAGCGGTTTGTGGCGG 5826
 Db 4917 CACTTATTTATACGAAACCGTCTGCGTAGCTCGGGTGGCATCCGATGATGTGG 4976
 QY 5827 AGCGGAACCTTATTTATCACTGCGCGGATCCGGTTTACCGGTTTACCGTGGCTCGG 5886
 Db 4977 AAGCGGAATCTTATTTATACCTCGCTGCAACAGGCTTGTGGTTTACCTTTGCGTTCA 5036
 QY 5887 CGTGGCCATCGTACGCTCGCGGCTTTATCGGAAATTTGGCGTCAAGTTAGACAG 5946
 Db 5037 CCGGGTCACTGCTGCTGCTATGATGATCTATCTGTAATCGGCAATTAAGTGAACAAG 5096
 QY 5947 CCGGTGTGCAAGCTGGGATTCAGCGGAGCGACCAAGTAAAGTTGAAAACTGGCGTGC 6006
 Db 5097 CAGGATAGCGGCTGGGATTCACACAAAGAACTAATCGTGGAGAGGTTACTTAT 5156
 QY 6007 TGGTGAAGGCGAGGTTTGGCGATCACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6066
 Db 5157 TAATGGCTGGCGCAATGAAGGCGCTTATATCTCGGTTGGCGGATAACAACGCTCGGATC 5216
 QY 6067 CTGGTCTGTGGCAGGCTCTCTGTTGA 6092
 Db 5217 CCAATTTGGCAGCGACCTCATTA 5242

RESULT 13

AAV40146/c

ID AAV40146 standard; DNA; 8625 BP.

AC AAV40146;

XX AAV40146;

DT 10-AUG-1999 (first entry)

XX Flavobacterium sp. R1534 nucleotide sequence.

DE Flavobacterium sp. R1534 nucleotide sequence.

KW Carotenoid; pigment; canthaxanthin; R1534; crtB; prephytoene synthase;

KW crtI; phytoene desaturase; crtY; lycopene cyclase; GGPP synthase; crtE;

KW crtW236; beta-carotene beta-oxygenase; food product; fermentation; ds.

XX Flavobacterium sp.

OS JP10155497-A.

XX 16-JUN-1998.

PD 16-JUN-1998.

XX

PF 02-DEC-1997; 97JP-00348653.
 XX 02-DEC-1996; 96EP-00810839.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA WPI; 1998-391048/34.
 DR P-FSDB; AAV69530, AAV69531, AAV69532, AAV69533, AAV69534.
 XX Preparation of carotenoid - comprises fermentation with transformed cell.
 XX Claim 1; Fig 7-21; 80pp; Japanese.
 XX The invention describes the preparation of carotenoid pigments e.g. canthaxanthins using a cell transformed by a vector having DNA sequences (a) to (e) or substantially homologous sequences. (a) a DNA sequence (crtE) coding GGPP synthase of Flavobacterium sp. R1534; (b) a DNA sequence (crtB) coding prephytoene synthase of Flavobacterium sp. R1534; (c) a DNA sequence (crtI) coding phytoene desaturase of Flavobacterium sp. R1534; (d) a DNA sequence (crtY) coding lycopene cyclase of Flavobacterium sp. R1534, and (e) a DNA sequence (crtW236) coding beta-carotene beta-oxygenase of a microbe E-396 (FERM BP-4283). The carotenoid or a carotenoid mixture can also be used in preparation of food products. The method is an improved method of fermentation for carotenoid production
 XX Sequence 8625 BP; 1458 A; 2898 C; 2964 G; 1295 T; 0 U; 10 Other;
 SQ
 Query Match 14.8%; Score 1037; DB 2; Length 8625;
 Best Local Similarity 56.6%; Pred. No. 1.2e-263;
 Matches 2038; Conservative 0; Mismatches 1535; Indels 30; Gaps 5;
 QY 2486 GTCCAGCAGCAAGTCGTGCTGGCGGAGCGACCTGATGCGCACCAATACGATGTGATT 2545
 Db 6984 GGGCGCGTCTGTAAGGACCGCGAAGGGCGGATCGCAATCATGAGCCATGATCTGCTG 6925
 QY 2546 TTGTCGCTGTGATGCGCAATGGCTTGTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2605
 Db 6924 ATCGCGGCGCGGGGCTGTCGGGTGCGTGTGCGGCTTTCGGGACCGCAGACCG 6865
 QY 2606 CAATGAAATGCTGTTCTGCGAGAGCGATGCGCATCCGCGAGGCAATCATACCTGGTGC 2665
 Db 6864 GATGCGCGCATGCTGATGCTGCGCGGCTGCGGCCCTCGGACGACGACACCTGGTCC 6805
 QY 2666 TTTTCATCACGCGATCTCAGCGCGGAAACAACTTCGCTGGCTGCAACCGCTGATACCGTG 2725
 Db 6804 TGCCACGACGCGATCTTTCGCCCGAATGGCTGGCGCGCTGTGCGCCATTCGTCGCGCG 6745
 QY 2726 CGTTGGTCAGGTTATCAGTGGTGTTCCTGGCTGCGCGCAATCTGGACGGGATTAAT 2785
 Db 6744 GAATGGACGCGATCAGGAGGTGCGGTTTCCCGACCAATTCGCGCGCTGACGACGCGTAT 6685
 QY 2786 TGTTCATCGCATCAGCGGATTTTCCCGCCCATCTTTACCGCGGATGGGTGACGATCTG 2845
 Db 6684 GGCTCGATCGAGGCGGCGCGCTGATCGGGTGTGCTGCA-----GGGTGCTGATCTG 6634
 QY 2846 TGGACAAACACAGCGGTAACAAGGTAAACCCACGAGGTGACCTGCGCGATGCGCGT 2905
 Db 6633 CGGTGGAATACGATGTCGCGACGCTGGACGATACCGCGCGGACGCTGACGCGGCTCG 6574
 QY 2906 GAATTCGTCGCAAGTGTGATGATGCTGCGGCTGCGCGGATGCGCGGATGCGGATCTG 2965
 Db 6573 CGGATCGAGGTTGCTGCTGATGACGCGCTGCTGCGGCTGCGGATGCGGATGCGGATGCG 6514
 QY 2966 CTGGGTTATCAGTGTGTTTCTGGACAGAGTGGAGCTGGCGGACCGCGCGCTGCGAG 3025
 Db 6513 GTGGGTTTCCAGAAATTCGTGGCGTGCAGATCGAGACGCGGCGGCGGCTGCGGCTG 6454
 QY 3026 CAGCGGATCTGATGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 3085
 Db 6453 CGCGGATGATGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 6394
 QY 3086 CTGCGGCTCAGCGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGG 3145

Db 6393 CTGCTTTCAGTCCACCGCATCTGATCAGGATACCGCTACAGCAGCGCGCGAT 6334
Qy 3146 CTGGCGGAGAACACCGCTGTCAGACATCCCGGATATGCAATCAGCAAGGCTGAGCG 3205
Db 6333 CTGACGATGCGCGCTGCGCAGCGCTCGCTGAGCATATGCGCGCAGCGCGGCTGAGCC 6274
Qy 3206 CTGAGTACGCTGCTGCGTGAAGACACCGGCATATACCGATTACCTTACCGTGAAGCGCAACATC 3265
Db 6273 GGGCAGGAGATGCGCGCGGGAAGGGGCATCTGCGCCATCGCGCTGGCCCATGACGCCATA 6214
Qy 3266 GATCGATTCTTGCAACACGACGCGCGCAAGGTGACGCGCTGCGCGCGCGCTGTTT 3325
Db 6213 GGTCTTGCGCGCACACCGCGCAGGGCGGTGCGCGTGGCTGGGGCAGGGCTGTTT 6154
Qy 3326 CATGCCACACCGGTTACTCTTGGCTGCGCGGTGCGCGTGGCTAGCGGAGTTGGTAGCAGCG 3385
Db 6153 CACCCCGTCACCGGATATTCGCTGCCCTATGCGCGCGCAGGTGCGGATGCGCATGCGCGCG 6094
Qy 3386 CTGTTGCCACCGATGCCCTACGCTCAGCTCAGCCAAACATATCGAACGCTTTGCCCGCTCAGCAG 3445
Db 6093 CCGGACCTGACGACCGCGCTCCCGCTGCGCGGTGCGGGCTGGGCCATCGATCGCGCG 6034
Qy 3446 TGGCGCGAACACGCAATTTTCGCTGCTGCTAAACCGCATGCTGTTTGGCGGTTAAGCGG 3505
Db 6033 GATCGCG--ACCGCTTCTCGCGGTGCTGAACCGGATGCTTCCGCGGTGCGCGGCC 5977
Qy 3506 CAGCAGCGCTGGCGGTGATGCAAGTTTACCGGCTCGATGCCGGGTAAATTAGCGCG 3565
Db 5976 GACCTGCGTATTCGCTGCGCGGTCTACCGGCTTCCGCGCTGCGCGACGCGCTGATCAGCGCG 5917
Qy 3566 TTTTACGCGCGCAACTGGCGCTGCGCGATAAACCGCGATTCTGTGGCGCAAGCGCGCG 3625
Db 5916 TTTATGCGCGCGCTGACATGCGCGACCGGCTTCGATCGTCAACCGGACCGCGCGCG 5857
Qy 3626 GTGCCATCGGTGAAGCGCTGCGCGGTGTTGAATTTCTGTGCAACCGAGGAGAAATA 3685
Db 5856 ATTCTGCTGCGAGCGCGCTGCGCTGCGCGAAACCGCGCTGCTGCGAGGAGAGCA 5797
Qy 3686 TGAAGCGCATTTATGATTGCGGAGGCTTTGGCGGCTGCGCTGCGGATTCGCTGCG 3745
Db 5796 TGAGTTCGCGCATCGTATCGCGGACAGGTTTGGCGGCTTGGCTTGGCATCGCGCTGCG 5737
Qy 3746 AAGCGCGGCGATACCAACCTTTACTGAGCAGCGCAAAACCGCGCGACCGGCT 3805
Db 5736 AATCGCGCGCATCGCGACCAACCATCGTAGCGCGCGCAAGCGCGCGCGCGCGCT 5677
Qy 3806 ATGTGTTGAGGACAGTGTCTTACCTTCGATGCGCGACCGCGGTGATCAGCATCCCA 3865
Db 5676 ATGTCTGGAACGATCAGGGCCACGCTTCTGATGCGAGGCGCGCGCTGCGACCGACCGG 5617
Qy 3866 GGGCCATCGAAGATTGTTTACGCTGGCAGGAAATTCGCTACGCGATTACGCTGAGCTGA 3925
Db 5616 ACAGCTCGGAGAGCTGTGGGCGCTCAGCGCGCAACCGATGAGCGTGAAGCTGCG 5557
Qy 3926 TGGCGGTAAACCGCTTCTATGCGCTGTGCGGAGATGGCAAAACAGCTTGATTAAGACA 3985
Db 5556 TGGCGGTCTGCGCTTCTACCGGCTGACATGGCGGACCGCGCGAGCTTCGAATACGTA 5497
Qy 3986 ATATCAGCGGTGCTGAGGACGACATCGCCACGTTCAATCGCAAGATGTAGAGGCT 4045
Db 5496 ACAGACGACGAGCTGATCCGCGAGGTGCGCTCTTCAATCCGCGCATGTCGATGCGT 5437
Qy 4046 ATCGTCAATTTCTGCTATTACGTTGAAGTATTTAGAGGGTATCTGAAACTCGCA 4105
Db 5436 ATCGCGCTTCCAGATTACCGCGAGGAGGTCTATCGAGGGGTATCTGAAGCTGGGA 5377
Qy 4106 CGGTGCGGTTTCTGAGGTGCGTGAATGCTGCGGCTGCGCGCGCAGGTTGGGACGCTGCG 4165
Db 5376 CACAGCGCTTCTTGAAGCTGGGCGAGATGCTGAACCGCGCGCGGCGCTGATGCGCGCTG 5317
Qy 4166 AAGCATGGCGAGGCTTACAGCATGTGGCGGAAATTTTACAGGACGATCATCTGCGTC 4225

Db 5316 AGGCATACCGCTCGGTCCACAGCATGGTGGCGCGCTTCTATCAGGACCCCGCATCTCGGCG 5257
Qy 4226 AGCGGTTTCTCTTCCACTCATTTGCTGGTGGCGGTAACTCTTTTTCGAAACGCTCATCGATCT 4285
Db 5256 AGGCTTCTCTGTTCCACACGCTGCTGGTGGCGGGAACCGGTTTTCGACCAGCTCGATCT 5197
Qy 4286 ATACCTTAATTCATGCGCTGAGCGGTGAATGGGGGCTGTGGTTTTCGCGCGGGCGGACCG 4345
Db 5196 ATGGCTGTATCCATCGCTGGAACCGCGCGCGCGCTCTGGTTTCGCAAGGGCGGACCA 5137
Qy 4346 GCGGCTGTGTCAGGGGATGCGCGACTGTTTCAGGAGACTTGGCGGCGAGCTGTACTGA 4405
Db 5136 ACCAGCTGGTGGCGGATGCTGCGCTGTTTCAGCGCTCTTGGCGGACGCTGCTGCTGA 5077
Qy 4406 ATGCCGAAGTGAAGCTGGAACACGAGCGGCAATCGCATTAGCGGCTTTCAGTTAGAGG 4465
Db 5076 ATGCCCGGTCACGCGGATCGACACGAGGGCGATGCGGCCACGCGGCTCACGCTGCTGG 5017
Qy 4466 GCGACGACGCTTCGATGCGCGCTGTGCTTCAATGCGGAGCTGGTGCATACCTACG 4525
Db 5016 ACGGCGGCGAGTTTGGCGCGGATACGCTGGCGGCAACGCGGAGCTGATGCACAGCTATC 4957
Qy 4526 ACAAACTGCTTCGCCACCATTCGCTGCAATGAACCTGCGACATCGCTGAAGCGTAAGC 4585
Db 4956 GCGACCTGCTGGGCAATACCGCGCGGGCGCACCAAGGCGCGATCTTGAACGCGCAGC 4897
Qy 4586 GCATGAGCAACTGCTGCTGTTTACTCTATTGTCCTGAAATCAGCCGATGAACAGCTCG 4645
Db 4896 GCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4837
Qy 4646 GCGACCAACGCTGCTGTTTGGCGCGGTTATCGTGAAGTGAATGAGATGAGATTTTCAACA 4705
Db 4836 CCAACCAACGCTGCTTTCGCGCGCGCTCAAGGGGCTGGTGAACGAGATCTTCAACG 4777
Qy 4706 GCAGCGAGCTGGCAGACGATTTTCACTTACCTGCAACGCGGCTTGCAGCAGCGATCCGT 4765
Db 4776 GCGACGCTGCGGACGATTTCTCGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4717
Qy 4766 CGCTGCAACCGCGCTGCGGAGCTTTTATGTTGTTAGCGCGGCTGCGGATCTCGGCA 4825
Db 4716 GCTTGGCGCGCGGAGTGTCCACGATTAACGCTTTCGCGCGCTTTCGCGATCTGGGCG 4657
Qy 4826 CCGCTGACATCGACTGGCAACAGGAAGACCGCGCTTTCGCGGATCGAATTTTCTGCTATC 4885
Db 4656 GCGCGATGCTGATTTGGAGAGCCGAGGCGCGGCTATGCGAGGCGATCTTCAGGAGAC 4597
Qy 4886 TGGAGCAGCATACATGCGCGGATTAAGTCAAGCAATTAAGTGAACACAGAAATTTTACGC 4945
Db 4596 TGGAGCGCGCGCATCCCGACCTGCGCAAGCACCTGACCGTCAAGCGCATCTTCAGCC 4537
Qy 4946 CGTTTGAATTTTCGACACGCTGATGCCATCAGGCTCGGCTTTCGCTGGAGCGCA 5005
Db 4536 CCGCGATTTTCAGCACCGAATGTCGCGCCATCAAGGCGGCTTTCGCTGCGAGCGCA 4477
Qy 5006 TTTTACGCAAGCGCTGCTTTCGCGCGATACCGGATGCGGATATCAGCAATCTCT 5065
Db 4476 TCCTGACCAATTCGCGCTTTCGCGCGATTAACCGGATCCGCGATCCGCAATCTCT 4417
Qy 5066 ATCTGCTGGTTCGCGTACGATCCAGCGCGGCGTCCCGGCGTGAATCGGTTGCGGCA 5125
Db 4416 ACATGCTGGGCGGCGACGATCCGCTGCGGCGATCCCGGCTGCTGCTGCGAGCGCA 4357
Qy 5126 AGGCACCGCGGCTGATGCTGGAGGATCGGCGGATCGCATGACAGCCGATCTTACTTGA 5185
Db 4356 AGGCACCGCGGCTGATGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 4297
Qy 5186 GCAAGTAAACGCAAAACATGCGGCTGGGCTGAAAGAGTTTTCGCGCACCGCGCGCAAGCTGT 5245
Db 4296 TTCCGAGG---GCCATCGCGAGGTTTCGCAAGCTTCGCGCAGGCGCGCAAGCTGAT 4240
Qy 5246 TGATGACCGGCGCGGACGACGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5305
Db 4239 GCGCGCGCGCATCCCGAGGATACGCTCATGCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 4180

QY 5306 TGTGATTGATGGGCAAAACCTCTGGCGAAGGCGGACGACAGCATGCGTGCAGAGACGGCA 5365
 Db |||||
 QY 4179 CGTGATCGACGGGCGAGGTGATGGTTCTGCCCGGAG-----GGGGGCGGCGACCA 4126
 Db |||||
 QY 5366 GGCAGTATGACGATCTGCAAAATGAACCCGCGCTACAGCGGCGGCACATGGA 5425
 Db |||||
 QY 4125 GCGCGGCTGGGGGCGCTCGCGCGGACACGCTGGCGCGCTGACAGAGACGGCCGAT 4066
 Db |||||
 QY 5426 TGAACCGGCTTTAGGGCGTTTTCAGGAAGTGGCGATCATTCACAGCTGCGCAACAAC 5485
 Db |||||
 QY 4065 GTCGCGCCCTTCGCGGCGCTCGCCAGGTGCGCGGCGCATGATTTCCCGGACCTTTG 4006
 Db |||||
 QY 5486 GCGGTTTGATCATCTGGAAGGCTTCTGATGATGACGACGACGACGACGACGACG 5545
 Db |||||
 QY 4005 GCCGATGGACCTGATCGAGGCTTTCGCGATGATGTCGCGATGCGGATACCGAGCCT 3946
 Db |||||
 QY 5546 CGATGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5605
 Db |||||
 QY 3945 GGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3886
 Db |||||
 QY 5606 CTAATGGCGCTGCGCGACGAGCGGTGCTGATCACGCTGCGATTTAGGACTGCGCTT 5665
 Db |||||
 QY 3885 GGTGATGGCGCTGCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3826
 Db |||||
 QY 5666 CCAGCTCACTAATGCGCGGACATTTGTAAGATGCGGAAATGCTGCTGCTGCTGCT 5725
 Db |||||
 QY 3825 CCAGCTGACAAATCGCTCGCGACGATGATGACGATGCGGACATGCGGCGCTGCTG 3766
 Db |||||
 QY 5726 GCGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5785
 Db |||||
 QY 3765 GCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3715
 Db |||||
 QY 5786 TCGTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5845
 Db |||||
 QY 3714 TTCGACGCGCTTATTCGCTGATCATCGCTGCTGACGCGCGGCGGCTTATGCG 3655
 Db |||||
 QY 5846 CTCGGCGGATCCGTTTACCGGTTTACCGTGGCTGCGGCTGCGGCGGCGGCGGCT 5905
 Db |||||
 QY 3654 CTCGGCGGCGGCGGCTTCCGATCTGCGCGCGGCTGCGGCTGCTGCTGCTGCTG 3595
 Db |||||
 QY 5906 TCGCGCGTTTATCCGAAATTTGGCTCAAGTTGACGACGCGGCTGCTGCTGCTG 5965
 Db |||||
 QY 3594 GCTGCTATCTATCGCGCAATCGGACGCGCATTCGCGAGGCTGCGCGGCGGCTG 3535
 Db |||||
 QY 5966 TTCGCGGACGCGCACGATGAAGGTGAAAGTTCGCTGCTGCTGCTGCTGCTGCTG 6025
 Db |||||
 QY 3534 CCAGCGGATCAGCAGTGAAGGCTGCAAGATCGGGCTTCTGGCGCGGAGGCTTGA 3475
 Db |||||
 QY 6026 GCGGATCACTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6085
 Db |||||
 QY 3474 CCGCGCGCATCGCGCTCGCGCGGCGGCAATCAGCCGCGGCGCTGCTGCAACCG 3415
 Db |||||
 QY 6086 TCG 6088
 Db |||||
 QY 3414 GCG 3412

RESULT 14

AAV40151/c

ID AAV40151 standard; DNA; 11233 BP.

XX

AC AAV40151;

XX

DT 10-AUG-1999 (first entry)

XX

DE DNA sequence of plasmid pZe4.

XX

KW Carotenoid; pigment; canthaxanthin; R1534; crtB; prephytoene synthase;

KW crtI; phytoene desaturase; crtY; lycopene cyclase; crtW2396; pZe4;

KW beta-carotene beta-oxygenase; food product; fermentation; ds.

XX

OS Synthetic.

XX JP10155497-A.
 XX 16-JUN-1998.
 XX 02-DEC-1997; 97JP-00348653.
 XX 02-DEC-1996; 96EP-00810839.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX WPI; 1998-391048/34.
 XX Preparation of carotenoid - comprises fermentation with transformed cell.
 XX Example 2; Fig 42-53; 80pp; Japanese.
 XX The invention describes the preparation of carotenoid pigments e.g. canthaxanthins using a cell transformed by a vector having DNA sequences (a) to (e) or substantially homologous sequences. (a) a DNA sequence (crtE) coding GPP synthase of Flavobacterium sp. R1534; (b) a DNA sequence (crtB) coding prephytoene synthase of Flavobacterium sp. R1534; (c) a DNA sequence (crtI) coding phytoene desaturase of Flavobacterium sp. R1534; (d) a DNA sequence (crtY) coding lycopene cyclase of Flavobacterium sp. R1534; and (e) a DNA sequence (crtW2396) coding beta-carotene beta-oxygenase of a microbe E-396 (FERM BP-4283). The carotenoid or a carotenoid mixture can also be used in preparation of food products. The method is an improved method of fermentation for carotenoid production
 XX Sequence 11233 BP; 2142 A; 3525 C; 3605 G; 1960 T; 0 U; 1 Other;
 SQ
 Query Match 14.8%; Score 1037; DB 2; Length 11233;
 Best Local Similarity 56.6%; Pred. No. 1.4e-263;
 Matches 2038; Conservative 0; Mismatches 1535; Indels 30; Gaps 5;
 QY 2486 GTGCAGCAGCAGTCTGCTGGCGAGCGGACCTGATGCGCACGCAATACGATGTGATT 2545
 Db |||||
 QY 7666 GGGCGCGTGTGAAGGACCGCGAAGGGGCGGATCGCAATACATGAGCCATGATCTGCTG 7607
 Db |||||
 QY 2546 TTGTCGCTGCTGGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2605
 Db |||||
 QY 7606 ATCGCGGCGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7547
 Db |||||
 QY 2606 CAATGAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2665
 Db |||||
 QY 7546 GATGCGCGATCGTATGCTGACGCGGCTCGGCGGCTCGGACGACGACACCTGGTCC 7487
 Db |||||
 QY 2666 TTTTCATCAGGATCTCAGCGCGCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2725
 Db |||||
 QY 7486 TGCCACGACACGATCTTTTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7427
 Db |||||
 QY 2726 CGTTCGTCAGGTTATCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2785
 Db |||||
 QY 7426 GATGACGCGATCAGGAGGTGCGGTTTCCGACCAATTCGCGCGCTGACGACAGGCTAT 7367
 Db |||||
 QY 2786 TGTTCATCGCATCAGCGATTTTTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2845
 Db |||||
 QY 7366 GGTTCGATCGAGGCGGCGCTGATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7316
 Db |||||
 QY 2846 TGGCAACACAGCGCTACACAGGTAACCCACGAGGTGACGCTGCGGATGGCGGT 2905
 Db |||||
 QY 7315 CGTGGGAATACGATGTCGCGACGCTGGAACGATACCGGCGGACGCTGACGCGGCTCG 7256
 Db |||||
 QY 2906 GAACTTGTGCGCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2965
 Db |||||
 QY 7255 CGGATCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7196
 Db |||||
 QY 2966 CTGGGTTATCAGGTTGTTTCTTGGCAAGAGTGGAGCTGGCGGACGCGCGGCTGCGAG 3025
 Db |||||
 QY 7195 GTGGGTTCCAGAAATTCGTGGCGCTCGAGATCGAGACCGACGCGCGGCTGCGAG 7136
 Db |||||
 QY 3026 CAGCCGATCTGATGATGCCACCGCTGATCAGCAAGGGGTTATGCTTTTCTTACAGC 3085

Db	7135	CGCCGATGATCATGACGACCGTTCGCGAGATGACGGGTACCGCTTCATCTATCTG	7076	6058	CCACGCCCTTCTGAAGCTGGGCCAGATGCTGAACGCCGCCGCGCTGATGCGCCTGC	5999
Qy	3086	CTGCCGCTCAGCGCGATCGGCTATTGATTGAAGTACCATTACGTTAAACACGCCCGCG	3145	4166	AAGCATGGCGGAGCGTCTACAGCATGGTGGGAAATTTATTACAGACGATCATCTGCGTC	4225
Db	7075	CTGCCCTTCAGTCCACCGCATCTGATCAGAGATACCGCTACAGGACGGCGGCGAT	7016	5998	AGGCATACCGCTCGGTCCAAGCATGGTGGCGCGCTTATCCAGGACCCGATCTGCGGC	5939
Qy	3146	CTGGCGGAGAACACCGCTCGTCAGCACATCGCCGACTATGCAATACAGCAAGGTGGAACG	3205	4226	AGSCGTTTTCTTCCATCATTTGCTGGTGGGCGGTAATCTTTTTCAAAGTCATCGATCT	4285
Db	7015	CTGAGCATGGCGCTGGCGAGCGGTCTGCTGGAATATGCCGAGCGGGCTGGACC	6956	5938	AGCCCTTCTGTTCCACACGCTGCTGGTGGCGGGAAACCCGTTTTTCGACCACTGATCT	5879
Qy	3206	CTGAGTACGCTGCTGGTGAAGACACGGCATATTACCGATTACCGTACGCGGCAATC	3265	4286	ATACCTTAATTCATGCGCTGGAGCGTGAATGGGCGTGTGGTTTTCCGCGCGGCGACCG	4345
Db	6955	GGGAGGAGATCGGCGCGAAAGGGGCATCTCTGCCCATCGCGCTGGCCCATGACGCCATA	6896	5878	ATGCGCTGATCCATGCTGGAAACGCGCGCGCGCGGTCTGGTTCGCAAGGGCGGACCA	5819
Qy	3266	GATCGATTCTGCAACAGCAGCGCGCAAGGTGTCAGCGGCTTCGCGCGCGGCTGTTTT	3325	4346	GCGCGTGGTGCAGGCGCATGGCGGACTCTTCGAGGACTTGGGCGCGGAGCTGTACTGA	4405
Db	6895	GGCTTCTGGCGGACACCGCGAGGGGGCGGTGCGGTTGGCTGGGGCAGGGCTGTTC	6836	5918	ACCAGCTGGTTCGGGGCATGTCGCCCTGTTTCGAGGCTTTCGGGCGCACGCTGCTGA	5759
Qy	3326	CATGCCACCAACCGGTTACTCTTTCGCTCCGCTGGCGGTAGCGGATGGTAGCAGCG	3385	4406	ATGCCGAAGTAGCCAGCTGGAAACACAGCGGCAATCGCATTTAGCGCGTTCAGTTAGAG	4465
Db	6835	CACCCGTCACCGGATATTGCTGCGCTATGCGCGGAGGTGCGGATGCCATCGCGCG	6776	5758	ATGCCCGCTCACCGGATCGACACCGAGGGGATCGCGCCACGCGGCGTCACTGCTGG	5699
Qy	3386	CTGTTGCCACCGATCCCTCAGCTCAGCCACATATCGAACGCTTTCGCCGTCAGCAG	3445	4466	GCGGACGACCTTCGATGCGCGCTGSCCTCCAAATGCGACGCTGTGCTATACCTACG	4525
Db	6775	CGCGACTGACGACCGCTCCGCGCGTCCGCGGCTGCGCGGTGGCCATCGATCGCGG	6716	5698	ACGGCGGCGGTTGCGCGCGGATACGCTGGCGGAGAAACGCGGACGTGACAGCTATC	5639
Qy	3446	TGCGCGGAACAGCGATTTCCTGCTCTGCTAAACCGCATCTGTTTTGGCGCGTAAGCG	3505	4526	ACAACTGCTTCGCCCATCTCGCTGGCAATGAAACGTCGCAATCGCTGAAGCTGAAGC	4585
Db	6715	GATCGCG--ACCGCTTCCTGGGCTGCTGMAACGGATGCTGTCGCGGCTGCCCGCC	6659	5638	GCGACTGCTGGGCCATACCGCGCGGGCGCACCAAGGCGCGGATCTCGAACCGCAGC	5579
Qy	3506	CAGACGCTGGCGGTGATGCAAGCTTTTACCGGCTCGATGCGCGGTAAATAGCGC	3565	4586	GCATGAGCAACTCGCTGTTGTTACTCTATTTTGGCTGAAATCAGCGCGCATGAACAGCTC	4645
Db	6658	GACCGTGCATCGCTGCTGACGCGTTCTACCGCTGCGCGGATGCTGTCGCGGCTGCC	6599	5578	GCTGGTGCATGTCGCTGTTTCGCTGCAATTCGCGCTTCTCAAGCGCCCGAGAACTT	5519
Qy	3566	TTTTACCGCGGCAACTCGCTCGCGATGAAACCGGATTCGTGCGGCAAGCGCGC	3625	4646	CGCACCAACACGCTGTTTTGGCGCGGTTATCGTGAAGTTGATGATGAGATTTCAACA	4705
Db	6598	TTCTATGCCCGGCGCTGACATTTGGCGGCGGCTGCGATGCTCACCGGACCGCGCC	6539	5518	CCACACACAGGTCATCTTCGCGCGCGCTACAAGGGCTGGTGAACAGATCTTCAACG	5459
Qy	3626	GTGCCCATCGGTGAAGCGCTGGCGGCTGTTGAATTCGTGCAACAGGAGAGAAAA	3685	4706	GCAGCCAGCTGGCAGAGATTTTTCATTTTACCTGTCAGCGCGCCTCGCAGCGGATCCGT	4765
Db	6538	ATTCGCTGTGCGAGCGCTGCGCTGCCCGAAGCGCCCTGCTGCGAGGAGAGCA	6479	5458	GGCCACGCTCGCGGAGATTTTCTCGATGATTCGCAATTCGCTCGCTGCGTGAACGATCC	5399
Qy	3686	TGAACCGCATTTGATGATTTGGCGAGCTTTGGCGCCTGGCGCTGGCGATTCGCTGC	3745	4766	CGCTGGCACCGCGCTGGCGAGCTTTTATGTTGTTAGCGCGCGTCCGCTCGCGCA	4825
Db	6478	TGATTCGCCCATCGTCAATCGCGCAGTTTCGCGGCTTGCGCTTGCCATCGCCTGC	6419	5398	GCCTGGCCCCCGAGGGGATGTCACGCAATTAAGTCTTGGCGCGCTTCGCACTCGGCGC	5339
Qy	3746	AAGCGGCGGCGATACCAACCACTTACTCGAGCAGCGCGCAAAACCGGGCGGACGCT	3805	4826	CGCTGCAATCAGTGGCAACAGGAAGAACCGCGCTTGGCGGATCGAATTTTCTTATC	4885
Db	6418	AATCGCGCGCATCGCGACCACTATCGTCAGGCGCCCGCAAGCCCGCGCGCGCT	6359	5338	GCGCCGATGTCGATTTGGGAAGCCGAGGCCCGGCTATGCCGAGCGCATCTTCAGGAAC	5279
Qy	3806	ATGTGTTGAGGACAGTGGCTTTTACCTTCGATGCGGACCCACGCTGATCACCGATCCCA	3865	4886	TGGAGCAGCATACATGCGCGGATTAAGTCAGCAATTAAGTGAACACAGATGTTTACGC	4945
Db	6358	ATGTCTGGAACGATCAGGCGCACGCTTCGATGAGGCGCCGACGCTGTCGACCCCG	6299	5278	TGGAGCGCGCGCCATCCCGGACCTTGGCGCAAGCACTGACCGTCAGCGCATCTTCAGCC	5219
Qy	3866	CGCCATCGAAGTGTTCACGCTGGCAGGAAATCGCTCAGCGATTAAGTGAAGCTGA	3925	4946	CGTTTGAATTTTCGCGACACGCTGCAATGCCATCAAGCTCGGCTTTCGCTGGAGCGCA	5005
Db	6298	ACAGCTCGAGAGCTGTGGGCGCTCAGCGGCCAACCGATGGAGCGTGAAGTGAAGCTG	6239	5218	CGCGCGATTTTCAGCACCGCACTGTCGCGCCATCAAGCGAGCGCTTTCGCTCGAGCGCA	5159
Qy	3926	TGCCGGTAAACCGCTTCATCGCTGCTGGGAGATGGAACACAGCTTGAATACGACA	3985	5006	TTTTGACGCAAGCGCTGTTTCGCGCGCATACCGGATGCGGATATCAGCAATCTCT	5065
Db	6238	TGCCGCTTCGCGCTTCTACCGGCTGACATGGCGGAGCGCGCGAGCTTCGATACGTGA	6179	5158	TCTGACGCAATTCGCGCTGTTCCCGCGCATAACCGGACCGCGGATCCGAACTTCT	5099
Qy	3986	ATAATCAGCGCTGCTGAGCAGAGATCGCGCATGTTCAATTCGCAAGATGTAGAAGCT	4045	5066	ATCTGTTGGTCCCGTACGCAATCAGGCGCGGGCTGCCCGGCTGATTCGCTCGGCA	5125
Db	6178	ACAGCAGCAGAGCTGATCGCGAGTCTGCTCTCTCTTCAATCCCGCGATGCTGATGCT	6119	5098	ACATCTGGGGGGGCGGACGCAATCCGGGTGGCGCATCCCGGTGCTCGTGGCAGCGCA	5039
Qy	4046	ATCGTCAATTTCTTGCCTATTACGTGAAGTATTTAGAGAGGTTTCTGAACTCGGCA	4105	5126	AGGCCACCGCAGGCTGATGCTGGAGGATCGCGCGCAATGATCCAGACCTTTACTTGA	5185
Db	6118	ATCGCGCTTCCACGATTAACCGCGAGAGGCTATTCGCGAGGGGTTCTGAACTGGGA	6059	5038	AGGCCACCGCGAGGTCATGCTGTGCGACCTGCGCGTGCATGACCGATCTCAGCGGAC	4979
Qy	4106	CGGTGCGGTTCTGAGAGTGTGATGATGCTGCGCGTCCGCGCGGAGTGGAGCTGTC	4165	5186	GCAAGTAAACGCAAAACCATGCGCGTGGGCTCGAAGAGTTTTCGCCACCGCGCGCAAGT	5245
				4978	TTCCGAGCG--GCCATCGCGAGGTTTCGCAAGCTTCGCGGAGGCGCGCAAGTGTAT	4922

DR P-PSDB; AAW06513, AAW06514, AAW06515, AAW06516, AAW00871, AAW06517,
DR AAW06518, AAW06519.
XX Flavobacterium gene sequences encoding carotenoid biosynthesis enzymes -
PT for the production of carotenoid(s), useful in foods and animal feeds.
XX Example 2; Fig 7; 80pp; English.
XX Genomic DNA (AAT45143) of Flavobacterium sp. R1534 includes genes of the
CC carotenoid biosynthesis pathway. The sequence was deduced from inserts of
CC 6 clones obtained from genomic libraries e.g. by PCR amplification (see also
CC AAT45144-45) and use of partial clones to screen the library. The
CC identities of the gene products (see also AAW06513-19 and AAW00871) were
CC detd. by examining carotenoid accumulation in E. coli hosts transformed
CC with deleted variants of the gene cluster. The isolated genes can be used
CC in different combinations to produce carotenoids in transformed host
CC cells. (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 8625 BP; 1458 A; 2898 C; 2964 G; 1295 T; 0 U; 10 Other;
Query Match 14.8%; Score 1035.4; DB 2; Length 8625;
Best Local Similarity 56.5%; Pred. No. 3.3e-263;
Matches 2037; Conservative 0; Mismatches 1536; Indels 30; Gaps 5;
QY 2486 GTGCCAGCAGCAAGTCGTGCTGGCGAGGGGACCTGATCGCGCACGCAATACGATGTGATT 2545
DB 6984 GGGCGGCGCTGCTGAAGGACCGCGAAGGGGCGGATCGCAATACATGAGCCATGATCTGCTG 6925
QY 2546 TTGGTGGTGTGCTGCACTGGCGAATGCTTGAATGGCGCTGCGTCTGCGTCAATTTGACGCCA 2605
DB 6924 ATCGGGGCGGGGCTGTGCGGTGCGTGTGATCGCGCTTGGCCGTTTCGCGACCGCAGACCG 6865
QY 2606 CAACCTGAATGCTGTGTTGGAGAGCGATGCGCATCCGGCAGGCAATCATACCTGTCG 2665
DB 6864 GATCGCGCATCGTGTGATCGATCGACGCGCGTCCGGCCCTCGGACGAGCACACCTGTGTC 6805
QY 2666 TTTTCATCAGCAGCATCTCAGCGCGCAACTTTCGCTGGCTGCAACCGCTGATACCGGTG 2725
DB 6804 TGCCACGACAGGATCTTTCGCCGATGCTGCGCGGCTGTGCGCCATTCGTGCGGC 6745
QY 2726 CGTTGGTCAAGTATCAGGTGCGTTTTCCTGCGCTGCGCGCAATCTGACGGGGATTA 2785
DB 6744 GAATGGACGATCAGGAGGTGCGGTTTCCCGACCAATTCGCGCGCTGACGACAGCTAT 6685
QY 2786 TGTTCCATCGCATCAGCGGATTTTGCCGCCATCTTTACGCGCGATGGGTGACGATCTG 2845
DB 6684 GGCTCGATCAGGCGGCGCGCTGATCGGGTGTGCA-----GGGTGTCATCTG 6634
QY 2846 TGGCAACAACAGCGTACACAGGTAAACCCACGACGATGACGCTGCGGATGGCGGT 2905
DB 6633 CGTTGGGAATACGATGTCGACGCTGGAAGATACCGCGCGACGCTGACGAGCGGCTCG 6574
QY 2906 GAACCTGTGCGCAAGTGTGATGTTGTTGCTGCGGCTGACGCGCCACATCTGCGAG 2965
DB 6573 CGGATCAGAGGTGCTGCTGATCGACGCGCTGTCGCGTGGTCGAGACCCCGACCTGACC 6514
QY 2966 CTGGTTATCAGTGTCTTTGGACAAGAGTGGCAGCTGGCGCAGCGCGCAGCGCTGCGAG 3025
DB 6513 GTGGGTTTCCAGAAATTCGTGGCGTTCGAGATCGAGCCGACGCCCCCATGGCGTCGAG 6454
QY 3026 CAGCCGATCCTGATGGATGCCACCGCTCGATCAGCAAGCGGGTATCGTTTGTCTACAG 3085
DB 6453 CGCCCGATGATCATGGACGCGACCGTTCCGAGATGGACGGGTACCGCTTCATCTATCTG 6394
QY 3086 CTGCGCTCAGCGCGGATGGCTATTGATTGAAGATACCCATTACGTTAACCGCCGCG 3145
DB 6393 CTGCGCTTCAGTCCACCGCGATCTGATCGAGGATACGCTACAGGACGCGGCGAT 6334
QY 3146 CTGGGGGGAACACACCGCTCTGTCAGCACATCGCCGATATGCAATACGAAAGGTGAGCG 3205
DB 6333 CTGGACGATGGCGCTGGCGCAGGCGTCTGCTGGACTATGCGCGCAGCGGGGTGGACC 6274
QY 3206 CTGAGTACCTGCTGCGTGAAGACGACGCGCATTTACCGATTACCTTGAGCGGCAACATC 3265

DB 6273 GGCAGGAGATCGGGCGGAAGGGGCATCTCTGCCATCGCGTGGCCCATGACGCCATA 6214
QY 3266 GATCGAATTTGGCAACAGCAGCGCGCCAGCGTGCAGGGCTCGCGCGGGCTGTGTT 3325
DB 6213 GGTTCGCGGACCAACGCGAGGGGGCGGTGCGCGTGGGTGGGGCAGGGGTGTC 6154
QY 3326 CATGCCACACCGGTTACTCTTTCCTGCGCTCGCGCGGTAGCGAGTGTGTAGACGCG 3385
DB 6153 CACCCGTCACCGGATATTGCTGCTGATCGCGCAGGTCCGATGCCATCGCGCG 6094
QY 3386 CTGTTGCCACCGAATCCCTCAGCTCAGCAACATATCAACGCTTTGCCGCTCAGCAG 3445
DB 6093 CGCGACTGACGACCGCTCGCGCGGTGCGCGGTGGGCAATCGATCGCGCG 6034
QY 3446 TGGCGGACAGCGATTTTTCGCTGCTAAACCGCATCTGTTTGGCCCGGTAAAGCGG 3505
DB 6033 GATCGCG--ACCGCTTCTTGGCGCTGCTGACCGGATGCTGTTCCGCGGCTGCCGCC 5977
QY 3506 CAGCAGCGCTGGCGCGTGTATGCAACGTTTACCAGCTCGATGCCCGGTTAATAGCCG 3565
DB 5976 GACGCTGCTATCGCTGCTGACGCGGTTCTACCGCTTCCGCGCAGCTGATCGAGCG 5917
QY 3566 TTTTACCGCGGCAACTGCGCTCGCGGATAAACCGGATTCGTGCGCAGCGCGCG 3625
DB 5916 TTCTATGCGGGCGCTGACATTTGGCGCACCGGCTTCGCACTGTCACCGGACCGCGCC 5857
QY 3626 GTGCCCATCGTGAAGCGCTGCGCGCTGTTGAATTTCTGTGAAACGAGGGAAGAAAA 3585
DB 5856 ATTCCGCTGTCGACGCGCTGCTGCTCCCGCAACGCGCCCTGTGTCAGGAGAGCA 5797
QY 3686 TGAACACGACTTATGTAATGGCGCAGCTTTGGCGCTGCGCTGCGGATTCGCTGC 3745
DB 5796 TGAGTTCCGCTATCGTATCGCGCAGGTTTCGCGGGCTTGGCTTGCATCGCCTGC 5737
QY 3746 AAGCGGGGCAATACCAACCACTTACTCGAGCAGCGCGCAAAACGCGGCGGACGCGCT 3805
DB 5736 AATCGCGCGGATTCGCGACCAACCATCGTCGAGGCGCGCAAGCGCGCGCGCT 5677
QY 3806 ATGTGTTTGAAGCAGTGGCTTTACTTCGATGCGGACCGACCGGATTTAGCTCGAGCTGA 3865
DB 5676 ATGCTGGAACGATCAGGCGCACGCTTTCGATGACGCGCCGACGCTGTCGACCGCCG 5617
QY 3866 GCGCATCGAAGTGTTCACGCTGCGCAGGAAATCGCTCAGCGATTTAGCTCGAGCTGA 3925
DB 5616 ACAGCTCGAGAGCTGTGGGCGCTCAGCGCGCAACCGATGGAGCGTGACGTGCGCTGC 5557
QY 3926 TGCCGCTAACGCGCTTCTATCGCTGCTGGGAAGATGGCAACACAGCTTGATACGACA 3985
DB 5556 TGCCGCTCTCGCGCTTCTACCGCTGACATGGCGGACGCGCGCAGCTTCGAATACGTGA 5497
QY 3986 ATAAATCAGCGCTGCTGAGCAGCAGATCGCCAGTTCAATCCGCAAGATGTAGAAGCT 4045
DB 5496 ACGACGACGAGCTGATCCGCGCAGGTGCGCTTCAATCCGCGGATTCGATGGCT 5437
QY 4046 ATCTGCAATTTCTGCTTATTCAGCTGAAGTATTTAGAGAGGTTTATCTGAAATCTCGGCA 4105
DB 5436 ATCGCGCTTCAACATTAACCGAGGAGTCTATCGCGAGGGTATCTGAAGCTGGGA 5377
QY 4106 CGGTGCGGTTTCTGAGGTGCTGATGCTGCGGCTCGCGCGCAGTTGGAGCTGTCG 4165
DB 5376 CCACGCGCTTCTGAAAGCTGGGCGCAGATGCTGAACGCGCGCGCGCTGATGCGCCCTGC 5317
QY 4166 AAGCATGGCGCAGCTTACAGCATGGTGGCGAAATTTATTTCAGGACGATCATCTGCGTC 4225
DB 5316 AGGCATACCGCTGCTGTCACAGCATGGTGGCGGCTTCATCCAGGACCGCATCTGCGGC 5257
QY 4226 AGGCGTTTCTTCCATCTATCTGTTGGGGCGGTAACTCTTTTGCACGTCATCGACT 4285
DB 5256 AGGCTTCTGTTTCCACACGCTGCTGGTGGCGGGAACCCGTTTTCGACCACTGATCT 5197
QY 4286 ATACCTTAATTCATGCGCTGAGCGTGAATGCGGGCTGTGGTTTTCGCGCGCGGACCG 4345

Db 5196 ATGGCTGATCATGGCTGGAAACGGCGCGGGCGTCTGGTTGCCAAGGGCGGCACCA 5137
Qy 4346 GCGCGCTGTGACGGGCATGGCGCGCATGTCTCAGGACTTGGCGCGCGAGCTGTACTGA 4405
Db 5136 ACCAGCTGTGCGGGCAGTGTGCGCCCTGTTCAGCGCTTGGCGGACGCTGCTGTA 5077
Qy 4406 ATGCGGAGTGAAGCAGCTGGAAACGAGCGGCAATCGCATTTAGCGGCGTTCAGTTAGAG 4465
Db 5076 ATGCCCGGTACCGGGATCGACACGAGGGCGATCGGCCACGGGCGTACGCTGCTGG 5017
Qy 4466 GGGGACGACGCTTCGATGCGCGCGTGTGGCCCTCAATGCCAGCTGGTGCACTACG 4525
Db 5016 ACGGGCGGAGTGGCGCGGATACGCTGGCCAGCAACGGCGACGTGATGACAGCTATC 4957
Qy 4526 ACAAACTGTTCGCCACCAATCCGCTGGCAATGAACGTGCGACATCGCTGAAGCGTAAGC 4585
Db 4956 GCGACCTGTGGGCCATACCCCGCGGCGCACCAAGCGCGCATCTCTGAACCGGCAGC 4897
Qy 4586 GCATGAGCAACTCGCTGTGTGACTCTATTTTGGCCTGNAATCAGCGGCATGACAGCTCG 4645
Db 4896 GCTGTCGATGTCGCTGTTCGTGCTGCAATTCGGCCTGTCCAAGCGCCCGGAGAACCTGG 4837
Qy 4646 CCGACACACCGCTCTGTTTGGCCCGCGTATCGTGAGTTGATCGATGAGATTTTCAACA 4705
Db 4836 CCCACACAGCTCATCTTCGGCCCGCGCTACAGGGGCTGTGAACGAGATCTTCAACG 4777
Qy 4706 GAGCGAGCTGCGACAGCAATTTTCACTTTACCTGCAACGCGCCCTGCGAGCAGCATCCGT 4765
Db 4776 GCGCAGCGCTGCGGAGCAATTTCTCGATGATCTGCAATTCGCGCCCTGCGGTGACCGATCCA 4717
Qy 4766 CGCTGGCACCGCCGCTGCGGAGCTTTATGTGTAGCGCGGTGCGGCATCTCGGA 4825
Db 4716 GCTTGGCCCCCGAGGGGATGTCACCGCAATACGTCCTTGGCGCCGTTCCGCATCTGGGCC 4657
Qy 4826 CGCTGACATCGACTGGCAACAGGAAGACCGCGCTTGGCGCATCGAATTTTGTCTATC 4885
Db 4656 GCGCGATGTGATTTGGGAGCGGAGGCCCGGCTATGCCAGCGCATCTTCGAGGAC 4597
Qy 4886 TGGAGCAGCACTACATGCGGGGATTAAGTACAGCAATTAAGTACACACAGAAATTTTACGC 4945
Db 4596 TGGAGCGCGCCCATCCCGACCTGGCGAAGCACCTGACCGTACAGCGCATCTTCAGCC 4537
Qy 4946 CGTTTGAATTTTCGACACGCTGCATGCCCATCAAGCTCGCGGTTTTCGCTGGAGCGCA 5005
Db 4536 CCGCCGATTTACAGCACCGCAATGTGCGGCCCATCACGGCAGCGCTTCTCGGTTCGAGCGCA 4477
Qy 5006 TTTTGACGCAAGCGCTGTTCCGCGCGCATAAACCGGATGCCGATACGCAATCTCT 5065
Db 4476 TCTGACGCAATCCGCTGTTCCGCGCGCATAAACCGGATGCCGATCCCGAATCTCT 4417
Qy 5066 ATCTGTTGGGTCCGGTACGATCCAGGCGCGGGGTGCGCGCGTGATCGGTTTCGGCCA 5125
Db 4416 ACATCGTGGGGCGGCGACGATCCGGGTGGGGCATCCCGGTGTGTTGGCAGCGCA 4357
Qy 5126 AGGCACCGCAGGCTGATGCTGGAGGATCGCGCGCAATGAATGACAGCCTTTACTTGA 5185
Db 4356 AGGCCACGCGCAGGTCATGCTGTGCGACCTGGCGGTGCGCATGACCGCATCTGACGCGGAC 4297
Qy 5186 GCAAGTAACGCAAAACCATGGCGGTGGCTCGAAGAGTTTTCGCCACCGCGCCCAAGCTGT 5245
Db 4296 TTCCGAGCG---GCCATCGCGAGGTTTCGCAAGCTTTCGCGAGCGGCGCAAGCTGAT 4240
Qy 5246 TGATGACGACGCGCGCAGCAGCTGATGCTGTATGCTGTGGTGTGCTCACTGGGATGA 5305
Db 4239 GCGCGCCGCGCATCCGCGAGGATACGTCATGCTCTATGCTGTGACGCGATCGCGATGA 4180
Qy 5306 TGTGATTTGGGCAAAACCTGGCGAAGCGGCGCAGCAGCATGCGTCAAGAGACCGCA 5365
Db 4179 CGTGATCGACGGGCGAGTGTGGGTTCTGCCCGCCGAG-----GCGGGCGGCGACCCCA 4126
Qy 5366 GGCAGTATGACGATCTGCAAAATGAACCCGCGCGCTACAGCGCGCGCATGGA 5425
Db 4125 GCGCGGCTGGGGGCGCTCGCGCGCGACACGCTGGCGCGCTGCGAGGACGGCCCGAT 4066

Qy 5426 TGAACCGCGTTTAGGGCGTTTACGAAAGTGGCGATCATTTACAGCTGCGCGCAACAAT 5485
Db 4065 GTGCGCGCCTTTCGCGCGCTCGCCAGGTGCGCGGGCGCATGATTTCCCGACCTTTG 4006
Qy 5486 GCGCTTTCATCATCTGGAAGGCTTCGCTATGATGATGACGCAACGAACTTACGCGGCTT 5545
Db 4005 GCGGATGACCTGATCGAGGGTTTCGCGATGATTCGCGGATTCGCGAATACCGAGCCT 3946
Qy 5546 CGATGACAGCTGCGCTTACTGCTACAGTTCGCGCGGTGTCGCTTTCGATGATGCGCGG 5605
Db 3945 GGATGACGTCTGGAATATTCCTACACGTGCGCGGGTTCGTTGGGCGTATGATGCGCGG 3886
Qy 5606 CGTAATGGGCTGCGCGACGAAAGCGGTGCTCATCAGCTTCGCAATTTAGATGCGCGTT 5665
Db 3885 GGTGATGGGCTGCGACGAGCATGCGGTGCTGATCGCGCTCGCATCTGGGCTTTGCGTT 3826
Qy 5666 CCAGCTCACTAACATTTGCGCGGCACATTTGTAGAAGATCCGAAATGTCGCTGCTATCT 5725
Db 3825 CCAGCTGACAAACATCGCTCGCGACGTATCGACATGCGGCCATCGGGCGCTGCTATCT 3766
Qy 5726 GCGCAATCTCGCTCGATCAGGCGGGATTTACGCGCGGATACGCTCACTGCAACCGCAACA 5785
Db 3765 GCCTCCGACTGGCTGGCGGAGCGG-----GGCGACGTTGAGGTCGCGTGCC 3715
Qy 5786 TCGTCAAGCTGCGCTCACTGGCAGCGGTTTGTGCGGAGGCGGAAACCTATTTATCA 5845
Db 3714 TTCGAGCGGCTTATTCGCTCATCATCGCTGCTTGAACGCGGCGAGCCCTATTTATGC 3655
Qy 5846 CTCGCGCGATCCGCTTACCGGTTTACCGGTTCGCTGCGCTCGGCTGGCCATCGCTACGGC 5905
Db 3654 CTCGCGCGGCGAGGGGCTTCGCGCATCTGCGCGCGCTGCGGCTGCTGATCGCGCGCG 3595
Qy 5906 TCGCGCGCTTTATCGCAAAATTTGGCGTCAAAAGTTTCAGCACGCGGTGTGACGCGCTGGGA 5965
Db 3594 GCTGCTATCTATCGCGCAATCGGACCGGCATCGGACGAGGTGCCCCGAGGCTATCG 3535
Qy 5966 TTCAGCGCAGCGCACAGTAAAGGTGAAAACTGGCGCTGCTGTTGTTGAAAGGGGCGAGTTT 6025
Db 3534 CCAGCGGATCAGACGTCGAAAGGCTGCCAAGATCGGGCTTCTGGCGCGCGAGGCTTGA 3475
Qy 6026 GCGCATCACTTCGCGTGTCTCGTTCCTGAACCGGCTCGGCTGCTGTTGCGAGCGTCC 6085
Db 3474 CCGCGCGCATCGCGCTTGGCGGCGGCAATCAGCGCGCGGCTGTTGACCCGAC 3415
Qy 6086 TCG 6088
Db 3414 GCG 3412

Search completed: January 25, 2005, 05:06:30
Job time : 2953 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 03:38:11 ; Search time 501 Seconds
(without alignments)
9929.763 Million cell updates/sec

Title: US-10-808-979-18
Perfect score: 6999
Sequence: 1 gcgatgcaatggaaatgt.....gacgttttagacatgcaata 6999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCUTS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2624.6	37.5	6918	1	US-07-783-705A-13
2	1035.4	14.8	8625	3	US-08-980-832-1
3	1035.4	14.8	8625	4	US-09-920-923B-1
4	1035.4	14.8	11233	3	US-08-980-832-27
5	1035.4	14.8	11233	4	US-09-920-923B-27
6	826.2	11.8	1479	1	US-07-783-705A-10
7	713.4	10.2	1518	1	US-08-095-726-7
8	713.4	10.2	1518	1	US-08-096-043-7
9	713.4	10.2	1518	1	US-08-096-623A-7
10	708.4	10.1	1522	1	US-08-095-726-9
11	708.4	10.1	1522	1	US-08-096-043-9
12	708.4	10.1	1522	1	US-08-096-623A-9
13	619.4	8.8	1482	3	US-08-660-645A-6
14	619.4	8.8	1482	3	US-09-298-718-6
15	619.4	8.8	1482	3	US-09-546-963-6
16	619.4	8.8	1482	4	US-09-547-267-6
17	515.6	7.4	1198	1	US-08-095-726-5
18	515.6	7.4	1198	1	US-08-096-043-5
19	515.6	7.4	1198	1	US-08-093-577-5
20	515.6	7.4	1198	1	US-08-096-623A-5
21	449.6	6.4	1083	1	US-08-331-004A-1
22	449.6	6.4	1083	5	PCT-US95-13937A-1
23	438	6.3	1235	1	US-08-095-726-13
24	438	6.3	1235	1	US-08-096-623A-13
25	430	6.1	1232	4	US-08-908-758-1
26	430	6.1	1232	4	US-09-023-587A-1
27	426.8	6.1	1235	1	US-08-095-726-15

28	426.8	6.1	1235	1	US-08-096-623A-15
29	419.8	6.0	1296	1	US-07-783-705A-8
30	414	5.9	891	1	US-07-783-705A-11
31	399.6	5.7	947	1	US-08-096-623A-16
32	399.6	5.7	947	1	US-08-096-623A-17
33	386.2	5.5	1149	1	US-07-783-705A-9
34	357.2	5.1	2886	1	US-08-663-310-7
35	357.2	5.1	2886	2	US-09-006-491-7
36	357.2	5.1	2886	3	US-09-335-919-7
37	346.2	4.9	909	1	US-07-783-705A-7
38	308.2	4.4	1157	1	US-08-095-726-3
39	308.2	4.4	1157	1	US-08-095-726-3
40	308.2	4.4	1157	1	US-08-096-043-1
41	308.2	4.4	1157	1	US-08-096-043-3
42	308.2	4.4	1157	1	US-08-093-577-1
43	308.2	4.4	1157	1	US-08-093-577-3
44	308.2	4.4	1157	1	US-08-096-623A-1
45	308.2	4.4	1157	1	US-08-096-623A-3

ALIGNMENTS

RESULT 1
US-07-783-705A-13
; Sequence 13, Application US/07783705A
; Patent No. 5429939
; GENERAL INFORMATION:
; APPLICANT: Misawa, No. 5429939ihiko
; APPLICANT: Kobayashi, Kazuo
; APPLICANT: Nakamura, Katsumi
; APPLICANT: Yamano, Shigeyuki
; TITLE OF INVENTION: DNA SEQUENCES USEFUL FOR THE
; SYNTHESIS OF CAROTENOIDS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladas & Parry
; STREET: 26 West 61 Street
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720Kb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: N/A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07783,705A
; FILING DATE: 19911023
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 1-103078
; FILING DATE: 21-APR-1989
; APPLICATION NUMBER: JP 2-53225
; FILING DATE: 05-MAR-1990
; APPLICATION NUMBER: US 07/519,011
; FILING DATE: 19-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwadron, Janet I.
; REGISTRATION NUMBER: 33,778
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-708-1935
; TELEFAX: 212-246-5959
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6918 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid (plasmid DNA)
; ORIGINAL SOURCE: Erwinia uredovora

Db 1618 CTCGGAAGCTTATGCGGCCAGTGAATAAATTTATGACTGGCTAATGCTGCTCATGACC 1677
Qy 1780 ACGTGATCCTCAAAACACGCGCGGCGTAAATTTGACGAGCGCGCGGATACATCAGT 1839
Db 1678 GTGTCATTGCGGAAACACAGCCACAGAAATGGGCTTAGCCGCCCGCAAAAGCTTCAACAGT 1737
Qy 1840 GCGTCTGCGCGTGCACAAATCAGCCAGATGGTGGCGGCTTGAATTTTCCAGCTCAGC 1899
Db 1738 GTTTTTCGCACTGCGCGCAATCAGCCAGCTTGTTCCTGAACTGGAATTTTCCCGCAAG 1797
Qy 1900 AACTGCGCCGCTGCTATCAGCGGTTGGGCGCACTCGCGGCCCGGTTTCTCTGCGCGCG 1959
Db 1798 CGTTACCGGCTGTTTTCATGCGGCTCGGCGCTCTGGCGAAACGACGACCGTCAACGT 1857
Qy 1960 TCCATGCGCCGCTG---GCCAGCGCTGCGTACGCGGCTGTTTATGCTCGCTCGGTAAGC 2016
Db 1858 CTTTCATCCGCTTATTTATCCTCAGAAAAAACC CGGATTTTTCGCTCGCTGGGCAAGC 1917
Qy 2017 TCGAAGGCGCATGCTTCCGCGCTTCTGCACTCTGGCGAGGCGTCCGCGAGCTCGCGC 2076
Db 1918 TTCAGGACACGCTTATGGGCTGTTTAAACGATAGTGAAGCCCTGTGAAGAAATTTGACG 1977
Qy 2077 TATCGCTGGTATCGCCCATTTGCGGGATTAAGCCGAACAGACGCAATCAGCTGGAGC 2136
Db 1978 GTCAGCTCTGTGTAGCCCACTGTGTGCTTTACGGACTCTCAGTGTGAAGAGCTGGCGC 2037
Qy 2137 TCGTGGCGGCGGTGGGTGACGGAATTTGTCATCAGCGCGCAGCCCTACAGCACGCGC 2196
Db 2038 GAAGCGCTCATACACAGGTGGTGGATTTTGGCCGATCAGTCAGCGCGGCTGTCTCAGGCGC 2097
Qy 2197 AGCTGTTTATCACTCATGCGCGGTTAAACAGCGCGCTGGAAGCACTGGAATCGGTACGC 2256
Db 2098 AGCTGGCGATCAACACGCGCGCATGAATACCGTACTGGAACGCAATTAATTACCGAAGC 2157
Qy 2257 CGATGCTGCGCTGCGGATTTGCTTTTGTATCAGCGCGGCTGGCGCGCATTTGATGGC 2316
Db 2158 CCTTTTACGCTTCGCTGGCTTTGATCAGCGCGGCGTCCGCTCAGCATCGTTTATC 2217
Qy 2317 ATGAGTTGGTCCCGCGCATCAGCTTTAGCGGTTTTCATCAACTGGAGCAGCATCTGC 2376
Db 2218 ACGGCATCGCAAGCGTCTCCGCTTTTACCACAGCCATGCTTTGGCTCGTCAAGTGC 2277
Qy 2377 AACAGCTGTACCGACGATGTTAGCGCTCAGGATGTCAGGATTCAGGCGAGCTGC 2436
Db 2278 GTTCATTTGTGACCAACGTCGATTTTACAGACGCAATGCGAATAATCCAGACGCGCTTC 2337
Qy 2437 AGCGCGCAGCGGTTGCCAGCGTGC CGCGACATGTCGACGAGCGGCTGTGCCAGCAGC 2496
Db 2338 GTTTGGCAGGGGCAACATGCGCGCTGCCGATATCATTTGAGCAGGTTATGTG-CAACGGT 2396
Qy 2497 AAGTCTGCTGGCGGAGGCGACCTGATGCGCAACGAATACGATGTGATTTTGGTCTGGTGC 2556
Db 2397 CAGCGCTGCTTAAAGTTGGGAGCGGCTATGCAACCGCATATGATCTGATTCCTCGTGGGCG 2456
Qy 2557 TCGACTGGCGAATGGCTTGAATTCGCTGGCTGCTGCTCAATTTGCGAGCCAACTGAAATG 2616
Db 2457 TGGACTCGCGAATGGCTTATCGCCCTGCTTTTACAGCAGCAACCTGATATCGGTAT 2516
Qy 2617 CCGTGTGCTGGAGAGCGATGCGCATCCGCGAGGCAATCATCTGCTGCTTTTTCATCACAG 2676
Db 2517 TTTGCTTATCGACGCGCAACCCAGCGCGGCGGGAATCATATGTTGTCATTTTCAACAGA 2576
Qy 2677 CGATCTCAGCGCGCAACAACTTTCGCTGCTGCAACCGCTGATTAACGCTGTTGGTCAAG 2736
Db 2577 TGAATTTGACTGAGAGCAACATCGTTGGATAGCTCGCTGGTGGTTTCATCACTGGCGCGA 2636
Qy 2737 TTATCAGGTGCTTTTTCGCTGCGCGCAATCTGGAACGGGATTAATGTTTCATCGC 2796
Db 2637 CTATCAGGTACGCTTTTCCACACGCGGCTGTAAGCTGAACAGCGGCTACTTTTGTATTAC 2696
Qy 2797 ATCAGCGAATTTTTCGCGCATCTTTTACGCGGATGGGTGAGCATCTGTGGACAAACAC 2856
Db 2697 TTCTCAGCGTTTCGCTGAGGTTTTCAGCGACAGTTTGGCCCGCACTTGTGGATGGATAC 2756

Qy 2857 AGCCGTAACAACAGGTAAACCCACGACGATGACGCTGGCGGATGCGCGTGAACATTGCTGC 2916
Db 2757 CGCGGTCGACAGGTTAATGCGGAATCTGTTCCGTTGAAAAAGGGTCAGGTTATCGGTGC 2816
Qy 2917 GCAAGTGTGATTTGATGCTCGCGGCTGCAGCGAGCGCACATCTGCAGCTGGTTATCA 2976
Db 2817 CGCGCGGTGATTTGACGCGGCGGTTATGCGGCAATTCAGCACTGAGGCTGGCTTCCA 2876
Qy 2977 GGTGTTTCTTGACAAAGATGTCAGCTGGCGCAGCGCACGCGCTGCAGCAGCCGATCCT 3036
Db 2877 GCGTTTATTTGCGCAGGAATGCGGATTTGAGCCACCGCATGTTTATCTCTCCCATAT 2936
Qy 3037 GATGATGCCACCGTCGATCAGCAAGCGGTTATCGTTTGTCTACACGCTGCGCTCAG 3096
Db 2937 CATGATGCCACGCTCGATCAGCAAAATGGTTATCGTTCGTGTACAGCTCGCGCTCTC 2996
Qy 3097 CGCCGATCGGCTTATGATTGAAGATACCATTCATTCAGTAAACAGCGCGCTCGCGGAGAA 3156
Db 2997 GCGGACCAAGATTGTTAATTTGAAGACACGCACTATATTGATAATGCACTATTAGATCCTGA 3056
Qy 3157 CACCGCTCTGTCAGCACATCGCGCACTATGCCAATCAGCAAGCTTGACGCTGAGTACGCT 3216
Db 3057 ATGCGCGCGCAAAATATTGCGACTATGCGCGCAACAGGCTTGGCAGCTTCAGACACT 3116
Qy 3217 GCTGCTGAAGACGACGCGCATATTACCGATTAACCTTGAGCGGCAACATCGATCGATTCTG 3276
Db 3117 GCTGCGAGAAGAACAGGCGCTTACCCTTACTCTGTGCGGCAATTCGCGACGCAATCTG 3176
Qy 3277 GCAACAGCAGCGCGCGCAAGCGTGCAGCGCTGCGCGCGCGGCTGTTTTCATGCCACAC 3336
Db 3177 G---CAGCAGCGCGCCCTGCGCTAGTGAATTAAGTGGCGCTGTTCCATCTTACCAC 3233
Qy 3337 CGGTTACTCTTTCGCGCTCGCGCTAGCGGATTTGTTAGCAGCGCTGTTGCCAC 3396
Db 3234 CGGCTTATCACTGCGCTGGCGTTGCGGTGCGCGCGCACTGAGTGCACTTGTATCTT 3293
Qy 3397 CGATGCCCTCAGCTCAGCGCAACATATCGAACGCTTTGCGCTGAGCAGTGGCGGCAACA 3456
Db 3294 TACGTCGCGCTCAATTCACCATGCAATTCAGCATTTTCCCGCGAGCGCTGGCAGCAGCA 3353
Qy 3457 GCGATTTTTCGCTCTGCTTAAACCGCATGCTGTTTTCGCGGTAAGCGCGCAGCAGCGCTG 3516
Db 3354 GGGCTTTTTCGCGATGCTGAATCGCATGCTGTTTTCGCGGACCGCGCGATTCAGCTG 3413
Qy 3517 GCGCGTGTATGCAACGTTTTTACCGCTCGATCCGCGTTAAATTAGCGCTTTTACCGCG 3576
Db 3414 GCGGTTATGACGCGTTTATGTTTACCTGAAGATTTAATTCGCGTTTTTATGCGGG 3473
Qy 3577 GCAACTGCGCTGCGCGATTAACCGGATTTCTGTGCGGCAAGCGCGCGTGCCTATCGG 3636
Db 3474 AAAACTCAGCTGACCGATCGGCTACGTAATTCGAGCGCAAGCGCGCTTTCGCGTATT 3533
Qy 3637 TGAAGCGCTGCGCGCTGTTGAAATTCGTCGAACAGGGAAGAAAAATGAACCGCACT 3696
Db 3534 AGCAGATTGCAAGCCATTTATGACACTCATCGTTAAAGAGCGACTATGAAGAACCACT 3593
Qy 3697 TATGATTTGGCGCAGGCTTTTGGCGGCTTGGCGCTGGCGATTTCGCTGCAAGCGCGGG 3756
Db 3594 ACGGTAATTTGGTCAGGCTTCGTTGCGCTGGCACTGGCAATTCGTTCTACAAGCTCGGG 3653
Qy 3757 ATACCAACACCTTATCTGAGCAGCGGCAAAACCGGCGGAGCGCGCTTATGTTGTAG 3816
Db 3654 ATCCCGCTTACTGCTTGAACAACGATGATAAACCGCGGCTCGGCTTATGTTCTACGAG 3713
Qy 3817 GACAGTGGCTTTTACCTTCGATGCGGACCCAGCTGATCACCGATCCCGCGCCATCGAA 3876
Db 3714 GATCAGGGGTTTACCTTTGATGACGCGCGACGGTTATCACCGATCCCGAGTCGCAATTGAA 3773
Qy 3877 GAGTTGTTTCACTGCGCAGGAAAAATCGCTCAGCGATTACGCTGAGCTGATGCGGTAACG 3936
Db 3774 GAACTGTTTGCACTGCGCAGGAAAAACAGTTAAAGAGTATGTCGAACTGCTGCGGTTACG 3833

Db 5976 GACCGTCGCTATCGCTGTCAGACGGTTCTACCGCTCGCCGACCGCTGATCGAGCGC 5917
Qy 3566 TTTTACGCGGGCAAACTGCGCTGCGGATATAACCGCGATTTCTGTGCGCAAGCCGCGC 3625
Db 5916 TTTATGCGGGCGCTGACATGTCGCGACCGCTTCGCATCGTCACCGGACGCGCGCC 5857
Qy 3626 GTGCCATCGGTGAAGCGCTGCGCGCTGTGTGAATCTGTGCAACAGGAGAAAGAAA 3685
Db 5856 ATTCCGCTGTGCGACGCGCTGCGCTGCGCCGAAACGCCCTCTGTCGAGGAGAGCA 5797
Qy 3686 TGAACGCACTTATGTGATTGGCGCAGGCTTTGGCGGCTGCGCTGCGGATTTGCGCTGC 3745
Db 5796 TGAGTTCCGCATCTCATCGCGCAGGTTTCGGCGGGCTTGCGCTTGCCATCGCCTGC 5737
Qy 3746 AAGCGCGGCGATACCAACCACTTACTGAGCAGGCGGCAAAACCGGGCGGACCGGCT 3805
Db 5736 AATCGCGCGCATCGCGACCACTCTGCGAGCGCGGACAAAGCCCGCGCGCGCGCT 5677
Qy 3806 ATGTGTTGAGCAGAGTGGCTTTACTTTCGATGCGGACCCACCGGTGATCACCGATCCA 3865
Db 5676 ATGTCTGGAAACGATCAGGGCCACGTTCTCGATGCGAGCCGACGCTGTCGACCGCCG 5617
Qy 3866 GGGCCATCGAAGTGTTCACGCTGGCAGGAAATCGCTCAGCGAATTAAGTCGAGCTGA 3925
Db 5616 ACAGCTGCGAGAGCTGTGGCGCTCAGCGGCCAACCGATGAGCGTGAAGCTGTC 5557
Qy 3926 TGGCGGTAAAGCGCTTCTATCGCTGTGCTGGAGAGTGGCAACAGCTTGATTAAGACA 3985
Db 5556 TGC CGCTCTCGCTTCTACCGGCTGACATGGGCGGACCGCGAGCTTCGAATACGTGA 5497
Qy 3986 ATAATCAGCGCTGTGAGCAGCAGATCGCCACGTTCAATCCGCAAGATGTAGAAGCT 4045
Db 5496 ACGACGACGAGCTGATCGCCAGTTCGCTCCTTCAATCCCGCGATGTCGATGCT 5437
Qy 4046 ATCGTCAATTTCTGCTATTCACGTAAGTATTTAGAGAGGTTATCTGAACCTCGCA 4105
Db 5436 ATCGCGCTTCCAGGATACGCGAGAGGTTCTATCGAGAGGTTATCTGAAGCTGGGA 5377
Qy 4106 CGGTGCGCTTTCGAGGTCGTAATGTCGCGCTGCGCGCGGATTTGGACGCTGTC 4165
Db 5376 CACGCGCTTCTGAAAGCTGGGCGAGATGCTGACCGCGCGCGCGCTGATGCGCTGC 5317
Qy 4166 AAGCATGGCGACGCTTACAGCATGCTGGCGAAATTTATTCAGGACGATCATCTGCTGC 4225
Db 5316 AGSCATACCGCTCGGTCACAGCATGCTGGCGGCTTCATCAGGACCGCGCATCTGCGC 5257
Qy 4226 AGCGTTTCTTCCATCTGCTGCTGGCGGTAATCTTTGCAAGCTCATGATCT 4285
Db 5256 AGGCTTCTGCTTCCACAGCTGCTGCTGGCGGGAAACCGCTTTTCGACAGCTCATCT 5197
Qy 4286 ATACCTTAATTCATGCGCTGAGCGTGAATGGGCGCTGTGTTTCCGCGCGCGGACCG 4345
Db 5196 ATGGCTGATCCATGCGCTGAAACGCGCGCGCGCTGCTGTTTCGCAAGGCGGACCA 5137
Qy 4346 GCGCGCTGTGAGGGATGCGCGGACTGTTTCAGGACTTGGCGCGGAGCTGTTACTGA 4405
Db 5136 ACCAGCTGTGCGGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5077
Qy 4406 ATCCGAAGTGAGCAGCTGGAACACAGCGCAATCGCATTTAGCGGCTTCAGTTAGAGG 4465
Db 5076 ATGCCCGCTCACCGGATCGACACGAGGGCGATCGCGCACGGGGCTCACGCTCTG 5017
Qy 4466 GCGGACGACGCTTTCGATGCGCGCGCTGTGGCTTCCAAATGCGGAGCTGTGCTACCTAG 4525
Db 5016 ACGGCGGCGAGTTGCGCGGAGTACGCTGCGGACGCGGAGCTGATGACAGCTATC 4957
Qy 4526 ACMACTGCTTGGCCACCATCGCTGCAATGAACGTCGACATCGCTGAAGCTAAGC 4585
Db 4956 GCGACCTGCTGGGCGATACCGCGCGCGGCGCACCAAGGCGGATCCTGAAACGGCAGC 4897
Qy 4586 GCATGACCACTCGCTGTTGTAATCTTAATTTGCGCTGAAATCAGCGCGCATGAACGCTG 4645
Db 4896 GCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4837

Qy 4646 CGCACACACCGCTCTGTTTGGCGCGGTTATCGTGAGTTGATCGATGAGATTTTCAACA 4705
Db 4836 CCACACACAGGCTATCTTCGGCGCGGCTTAAGGGCTGTGTAACAGATCTTCAACG 4777
Qy 4706 GCAGCCAGCTGGCAGACGATTTTCTACTTTACTGCAAGCGCGCTGAGAGGATCCGT 4765
Db 4776 GGCACGCTTCCGCGAGATTTCTCGATGTATCTGATTTCTGCTGCTGCTGCTGCTGCT 4717
Qy 4766 CGCTGGACCGCGCTGCGGCTGCGGCTTTTATGTTAGCGCGGCTGCGCATCTGGCA 4825
Db 4716 GCCTGGCGCGCGAGGAGATGTCACGCAATTAAGTCTTGGCGCGCTTCCGATCTGGCC 4657
Qy 4826 CCGCTGATCATGCTGCAACAGAGAGACCGGCTTTGCGCGATCGAATTTTCTTATC 4885
Db 4656 GCGCGATGTCGATTGGAGCGGAGCGCGCTGATGCGAGCGCATCTTCAGGAAC 4597
Qy 4886 TGGAGCAGCACTACATCGCGGATTAAGTACGCAATTTAGTACACACAGATGTTTACGC 4945
Db 4596 TGGAGCGCGCGCATCCCGACCTGCGCAACCTGACCGCTGACCGCATCTTCAGCC 4537
Qy 4946 GCTTTGATTTTCGCGACACGCTGATCCCATCACGCTCGGCGTTTTCGCTGAGCGCA 5005
Db 4536 CCGCGGATTTTACGACCGAATCTGTCGCGCATCACGCGAGCGCTTCTCGTTCAGCGCA 4477
Qy 5006 TTTTGAAGCAAGCGCTGCTTCCGCGCATAAACCGCGATGCGGATTCAGCAATCTCT 5065
Db 4476 TCCTGACGAATCCGCTGTTCCGCGCATAAACCGCGAGCGGATCCCGAATCTCT 4417
Qy 5066 ATCTGTTGGTTCGCGTACGATCCAGCGCGGCGTCCCGCGGCTGATCGTTTCGCGCA 5125
Db 4416 ACATGTTGGGCGGCGACGATCCGCGTGGGCGATCCCGGCTGCTGTTGGCAGCGCA 4357
Qy 5126 AGGCACCGCGCATGCTGAGGATCGCGCGCAATGAATCGACAGCGCTTTACTGA 5185
Db 4356 AGGCACCGCGCATGCTGCTGCGACCTGCGCGCTGCGATGACCGATCTGACGGGAC 4297
Qy 5186 GCAAGTAAACCAATGCGGCTGCGCTGCAAGATTTCCGACCGCGCGGCAAGCTGTT 5245
Db 4296 TTCCGAAGCG--GCCATCGCGAGGTTTCCAAAGCTTCGCGAGGCGGCGCAAGCTGAT 4240
Qy 5246 TGATGACCGCGCGCGCATGCTGATGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTG 5305
Db 4239 GCGCGCGCGCATCCGAGGATACGCTCATGCTCTATGCTGCTGCTGCTGCTGCTGCTG 4180
Qy 5306 TGTGATGATGGGCAACGCTGGGCGAAGCGCGACGCGATGCGCTGCGGAGCGCGCA 5365
Db 4179 CGTGATCGACGGCGAGGTGATGGGTTCTGCCCCGAG-----GCGGCGGCGGACCCACA 4126
Qy 5366 GGCAGTATGCAATCTGCAATTTGAACCGCGCGCTTACAGCGCGCGGCAATGGA 5425
Db 4125 GCGCGGCTGGGCGGCTGCGCGCGCACGCTGGCGCGCTGCGAGGAGCGGCGCGAT 4066
Qy 5426 TGAACCGGCTTTAGGCGGTTTTCAGGAAAGTGGCGATCATTCACAGCTGCGCGCAACT 5485
Db 4065 GTCCGCGCTTTCGCGCGCTGCGCGAGTGGCGCGGCGATGATTTTCCGCGACCTTTG 4006
Qy 5486 GCGCTTGTGATCATCTGGAAGGCTTTCGCTATGATGTCACGCAACGAATTAAGCGAGCT 5545
Db 4005 GCGGATGACCTGATCGAGGTTTTCGCGATGATGATGATGATGATGATGATGATGATGAT 3946
Qy 5546 CGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5605
Db 3945 GATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3886
Qy 5606 CGTAATGGCGCTGCGGACGAAAGCGGCTGCTGATCAGCGCTGCGATTTAGGATGCGCT 5665
Db 3885 GGTGATGGCGCTGCGGACGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 3826
Qy 5666 CAGCTCATAACATGCGCGGACATTTGTAAGATGCGGAAATGCTGCTGCTGCTGCTGCT 5725
Db 3825 CCAGCTGACGAACATCGCTGCGGACGATGATGATGATGATGATGATGATGATGATGATGCT 3766

Qy	5726	GC	CGCAATCTTGGCTCGATCAGGCGGGATTAACGGCCGATACGCTGATCGACCGCAACA	5785
Db	3765	GC	CTGCCGAGCTGGCTGGCCGAGCGGG-----GGCAGCGTTTGAAGGCTCCGGTGCC	3715
Qy	5786	TC	TGCAGCGCTCGCCCTCACTGGCAGCGCTTTAGTCGGCGAGCGGAAACCTATTATTCA	5845
Db	3714	TT	TGCGAGCGCTCTATTCCGTCAATCCCGCTGCTTGACGGCGCCGAGCCCTATTATGC	3655
Qy	5846	CT	CGCGCGCATCCGGTTTACCGGGTTTACCGCTGCGCTCGCGCTGGGCCATCGCTACGCG	5905
Db	3654	CT	CGCGCGCGCAGGGCTTCGCGCATCTGCCCGCGCTGCGCGTGGTTCGATCGCCGCCG	3595
Qy	5906	TC	GGGGGGTTTATCGCGAAATTGGCTCAAAATTGACGACCGCCGGTGTGCACGCCCTGGGA	5965
Db	3594	GCT	CGCTATCTATCGCGCAATCGGGACGCGCATCCGCGAGGCTGGCCCCGAGGCTATCG	3535
Qy	5966	TT	CACGCGCAGCGCACAGTAAAGTGAAGAACTGGCGCTCTCGTGAAGAGGGCAGGTTT	6025
Db	3534	CC	AGCGGATCAGCACATCGAAGGCTGCCAAGATCGGGCTTCTGCGCGCGGAGGCTTGGGA	3475
Qy	6026	GG	CGCATCACATTCCGGTGTGTCTCGCTCTGAACCGCGCTCCGCTGGTCTGTGGGACGCGTCC	6085
Db	3474	CG	CGGCGCGATCGCGCTTGGCGCGCGGCGGAAATCAGCCGCGACGGCTGTGGACCGGACC	3415
Qy	6086	TC	G 6088	
Db	3414	GCG	3412	

RESULT 3

US-09-920-923B-1/c

; Sequence 1, Application US/09920923B

; Patent No. 6677134

GENERAL INFORMATION:

APPLICANT: Pasamontes, Luis

APPLICANT: Tsygankov, Yuri

TITLE OF INVENTION: Fermentative Carotenoid Production

FILE REFERENCE: 15464 US (C38435/125944)

FILE REFERENCE: 13404 03 (C30433/123344)
: CURRENT APPLICATION NUMBER: US/09/920.9231

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

/ CURRENT FILING DATE: 2001-06-02
 : PRIOR APPLICATION NUMBER: 08/980,832

; PRIOR AFFILIATION NUMBER: 08/
 : PRIOR FILING DATE: 1997-12-01

; PRIOR FILING DATE: 1997-
 : NUMBER OF SEQ ID NOS: 66

```

; NUMBER OF SEQ ID NOS: 66
: SOFTWARE: PatentIn version 3.1

```

```

; SOFWAKE: F
; SEQ ID NO 1

```

```

; SEQ ID NO 1
:
: ENGINE: 8625

```

LENGTH: 80
TYPE: DNA

TYPE: DNA
ORGANISM: Escherichia coli P1534

ORGANISM
FEEDING:

; FEATURE: ~~NAME / VERB~~ ~~NAME / VERB~~

```
; NAME/KEY: misc_feature
```

; LOCATION: (8348)..(8349)

OTHER IN:

; FEATURE:

```
; NAME/KEY: misc_feature
```

; LOCATION: (8539) .. (8540)

; OTHER IN:

FEATURE:

```
; NAME/KEY: misc_feature
```

; LOCATION: (8581).. (8581)

OTHER IN:

FEATURE:

```

; NAME/KEY: misc feature

```

NAME/UNIT: MISS LOCATE
LOCATION: (8590) .. (8590)

OTHER IN

: FEATURE:

```

FEATURE:
NAME/KEY: misc feature

```

```
; NAME/KEY: MISC_LEADUE
: LOCATION: (8592) (8592)
```

LOCATION :
OTHER :
:

OTHER INFORMATION: UNSUB.
FEATURE:

```

; FEATURE:
; NAME/REV. misc feature

```

```

; NAME/KEY: misc_feature
; LOCATION: (8C03) (8C04)

```

; LOCATION: (8 OTHER TOWN)

3506 CAGACGCGTGGCGCTGATCAACGTTTACCGGCTCGATGCGGGTTAATTAGCGCG 3565
 5976 GACCGTCTATCGCTGCTGAGCGGTTCTACCGCTCGCGCAGACCGGCTGATCGAGCGC 5917
 3566 TTTTACGCGGCGAACTGCGGCTGGCGATTAACCGGATTTCTGCGGCAAGCGCGCG 3625
 5916 TTTTATGCGGCGGCTGACATGCGGCTGCGCATGCTACCGGACCGCGCGCGCG 5857
 3626 GTGCGCATCGTGAAGCGCTGCGCGCTGTTGAATTTCTGTCGAACACGAGGAAGAAAA 3685
 5856 ATTTCGCTGTCGAGCGCTGCGCTGCTGCCGGAACCGCCCTGCTGACGAGAGAGCA 5797
 3686 TGAAGCGCATTTATGTGATGCGCAGGCTTTGGCGGCTGGCGCTGCGGATTTGCGCTGC 3745
 5796 TGAATTCGCCCATCGTCAATGCGCGAGGTTTCGGCGGCTTTCGCTTCCCATCGCGCTGC 5737
 3746 AAGCGCGGCGATACCAACCACTTACTGACGACGCGACAAACCGCGCGGAGCGCGCT 3805
 5736 AATCGCGCGGCGATCGGACCAACATCGTCGAGCGCGGACGAGCCCGCGCGCGCT 5677
 3806 ATGTGTTTGAAGACAGTGGCTTTTACCTTCGATGCGCGACCCACGCTGATACCGATCCCA 3865
 5676 ATGTCTGGAACGATCAGGCGCACGCTCTTCGATGCGAGCCCGACGCTGCGACCGCG 5617
 3866 GCGCCATCGAAGGTTGTTACGCTGGCAGGAATCGCTCAGCGATACGTCGAGCTGA 3925
 5616 ACAGCTGCGAGAGCTGTGGCGCTCAGCGGCGCAACCGATGAGCGTGACGTCGATGGCT 5557
 3926 TGCCTGTAACCGCTTCTATCGCTGTGCTGGGAAGATGGCAACAGCTTTGATTAAGCA 3985
 5556 TGCCTGCTCGCTTCTTACCGCTGACATGGCGGAGCGGCGCGAGCTTCGAATACGTCGA 5497
 3986 ATAATCAGCGCTGCTGAGCAGCAGATGCCAGCTTCAATCGGCAAGATGAGAGGCT 4045
 5496 ACAGCAGCAGAGCTGATCGCGCAGTTCGCTTCAATCGCGCGGCTGATGCGCTGC 5437
 4046 ATGCTCAATTTCTGCTATTTACGTTGAAGTATTTAGAGAGGTTTATCTGAACCTCGGCA 4105
 5436 ATGCGCGCTTCCAGATTACCGCAGGAGTCTATCGGAGGGTATCTGAGCTGGGA 5377
 4106 CGGTGCGCTTCTGAGGCTGCTGACATGCTGCGGCTGCGCGCGGCTGATGCGCTGC 4165
 5376 CCACGCTTCTTGAAGCTGGCGCAGATGCTGAACCGCGCGGCTGATGCGCTGC 5317
 4166 AAGCATGCGCAGGCTTACAGATGCTGGGAAATTTATTCAGGACGATCATCTGCTGC 4225
 5316 AGGCATACCGCTCGCTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5257
 4226 AGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4285
 5256 AGGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5197
 4286 ATACCTTAATTCATGCGCTGAGCGTGAATGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 4345
 5196 ATGCGCTGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5137
 4346 GCGCGCTGCTGCGGCGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4405
 5136 ACCAGCTGCTGCGGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5077
 4406 ATGCGGAAATGAGCGCTGGAACACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4465
 5076 ATGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5017
 4466 GCGGAGCGCTTCTGATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4525
 5016 ACGGCGGCGAGTTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4957
 4526 ACAAATGCTTCTGCGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4585
 4956 GCGACTGCTGCGGCGATACCGCGCGGCGCACCAAGCGCGGCTGCTGCTGCTGCTGCTGCT 4897
 4586 GCATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4645

4896 GCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4837
 4646 CCGACCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4705
 4836 CCGACCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4777
 4706 GCGACGCTGCGCAGCAGATTTTCACTTTCATCTGACGCGCTGCTGACGAGCAGTTCGT 4765
 4776 GCGACGCTGCGGAGCAGATTTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4717
 4766 CCGTGCACCGCGCGCTGCGGAGCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4825
 4716 GCTGCGCGCGGAGGATGCTCAGCATTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4657
 4826 CCGTGCATTCGACTGCGCAACAGGAGGACCGGCTTGGCGGATCGAATTTTGTCTATTC 4885
 4656 GCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4597
 4886 TGGAGCAGCATCTACATGCGCGGATTTAGTCAGCAATAGTGACACACAGATGTTTACGC 4945
 4596 TGGAGCGCGCGCATCCCGACCTGCGCAAGCACCTGACCGCTGACCGCGCATCTTCAGCG 4537
 4946 CGTGTTGATTTTTCGCGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5005
 4536 CCGCGGATTTTTCAGCACCGAATGCTGCGGCTGCTGCGGAGCGCTTCTCGTTCGAGCGCA 4477
 5006 TTTTACGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5065
 4476 TCTGAGCGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4417
 5066 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5125
 4416 ACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4357
 5126 AGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5185
 4356 AGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4297
 5186 GCAAGTAAACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5245
 4296 TTTGCGAGCG--GCCATCGCGAGGTTGCGAAGCTTTCGCGAGCGCGCGCTGCTGCTGCTGCTGCT 4240
 5246 TGATGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5305
 4239 GCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4180
 5306 TGTGATTTGATGCGGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5365
 4179 CGTGATCGACGCGCGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4126
 5366 GCGACGATGCGAGCATCTGCAAAATTGAAACCGCGCGCTTACAGCGCGCGCGCACATGGA 5425
 4125 GCGCGCGCTGCGGCGCTGCGCGCGCACACGCTGCGCGCGCTGCGACGAGGAGCGCGCGAT 4066
 5426 TGAACCGCGCTTACGCGGCTTTCAGGAAGTGGCGCATTCATCCAGCTGCTGCGCGCAAACT 5485
 4065 GTGCGCGCTTTCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTTTCGCGGCTTTCG 4006
 5486 GCGCTTTCATCTGGAAGGCTTTCGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5545
 4005 GCGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3946
 5546 CCGTGCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5605
 3945 GCGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3886
 5606 CCGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5665
 3885 GGTGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3826
 5666 CCGCTCCTTACATTCGCGCGCATTTGAGAGATGCGGAAATGCTGCTGCTGCTGCTGCTGCTGCT 5725

Db 3825 CCAGCTGACGAACATCGCTCGCGAGTGAATGACAGATCGGCCATCGGGCGCTGCTATCT 3766
Qy 5726 GCCGCAATCTCGCTCGATCAGCGGGGATTAACGCGCCGATACGCTGACTGCAACGCAACA 5785
Db 3765 GCCTGCCGACTGGCTGGCCGAGCGGG-----GGCGACGGTTGAGGGTCCGGTGCC 3715
Qy 5786 TCGTGCAGCGCTCGCTCACTGGCAGCGGTTAGTGGCGGAGGCGGAACCTATTATCA 5845
Db 3714 TTGCGACGCGCTCTATTCGGTCAATCCGCTGCTTGACGCGCGCGAGCCCTATTATGC 3655
Qy 5846 CTCGCGCGCATCCGGTTTACCGGGTTACCGGTTTACCGCTCGGCTGGGCTGATCGTACGCG 5905
Db 3654 CTCGCGCGCGAGGGCTTCGCACTGCGCGCGGCTCGGCTGGTGGTCAATCGCGCGCC 3595
Qy 5906 TCGCGCGGTTTATCGGAAATTTGGCGTCAAGTTTCAACGCGCTGGTGTGACGCGCTGGA 5965
Db 3594 GCTGGTATCTATCGCGCAATCGGACGCGCATCGGACGGGTTCGCGCGCGGCGGCTATCG 3535
Qy 5966 TTCACGGCAGCGCACAGTAAAGGTGAAAACTGGCGCTGCTGGTGAAGGGGCGAGTTT 6025
Db 3534 CCAGCGGATCAGCAGCTCGAAGGCTGCCAAGATCGGGCTTCTGGCGCGCGAGGCTTGA 3475
Qy 6026 GCGCATCACTTCGCGTGTCTCTGCTCAACCGGCTCGGCTGGTGTCTGTCAGCGTCC 6085
Db 3474 CCGCGCGCATCGCGCTCGCGCGCGCGGCGGAATCAGCCGCGAGGCTGTGACCGGACC 3415
Qy 6086 TCG 6088
Db 3414 GCG 3412

RESULT 4

US-08-980-832-27/c
; Sequence 27, Application US/08980832B
; Patent No. 6291204
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; FILE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/08/980,832B
; CURRENT FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 27
; LENGTH: 11233
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Plasmid pZea4
US-08-980-832-27

Query Match 14.8%; Score 1035.4; DB 3; Length 11233;
Best Local Similarity 56.5%; Pred. No. 3.1e-273;
Matches 2037; Conservative 0; Mismatches 1536; Indels 30; Gaps 5;
Qy 2486 GTGCCAGCAAGTCTGCTGCGGAGCGGACCTGATCGCACGCAATACGATGTGATT 2545
Db 7666 GGGCGCGCTGTAAGGACCGGAGGGCGGATGCAATACATGAGCCATGATCTGCTG 7607
Qy 2546 TTGGTCGGTGTGACTGCGGAATGGCTTGATTTGGCTGCTGCTGCTGCTGCTGCTG 2605
Db 7606 ATCGGGGCGGGGCTGTCGGTGTGCTGATCGCGCTTGGCGCTTGGCGACCGGACG 7547
Qy 2606 CAACTGAATGCTGTTGCTGGAGAGCGATGCGCATCGCGGAGGCAATACATACCTGTCG 2665
Db 7546 GATGCGCGCATGCTGATGCTGACGCGCGGCTCGGCGCTTGGCGACCGACCTGCTGTC 7487
Qy 2666 TTTCATCACAGGATCTAGCGCGCGGAACTTTCGCTGGCTGCAACCGCTGATTACCGTG 2725
Db 7486 TGCCACGACAGGATCTTTCCGCCGAATGGCTGGCGCGCTGTGCGCCATTCGTGCGGCG 7427
Qy 2726 CGTTGGTCAAGTTATCAGGTGGCTTTTCTGCTGCGCTCGCGCGCAATCTGACCGGGATTAT 2785

Db 7426 GAATGGAATCGATCAGAGGTGCGTTTCCGACCAATTCGCGCGCTGACGACAGGCTAT 7367
Qy 2786 TGTTCATCGCATCAGGGGATTTTCCCGCCCATCTTTACGCGCGATGGGTGACGATCTG 2845
Db 7366 GGTTCGATCAGGGCGGCGGCTGATCGGGCTGCTGCA-----GGGTGCTGATCTG 7316
Qy 2846 TGGCAAAACACAGCGCTCAACAGGTAAACCCACGAGGTGACGCTGCGGATGCGCGT 2905
Db 7315 CCGTGGAAATACGATGTCGCGACGCTGAGACGATACCGGCGGACGCTGACGAGCGCTG 7256
Qy 2906 GAACTTGTGCGCAAGTGGTGAATGATGTCGCGGCTGCGACCGGACGACATCTGCGAG 2965
Db 7255 CCGATCGAGGCTGCTGCGTGATGACGCCCTGCTGGCTCGAGACCCCGACCTGACCC 7196
Qy 2966 CTGGGTTTATCAGTGTCTTCTTGCAACAAGAGTGGCAGCTGGCGCAGCGCACGCGCTG 3025
Db 7195 GTGGGTTTCCAGAAATTCGTGGCGCTCGAGATCGAGACCGACCGCCCTGCGGCTG 7136
Qy 3026 CAGCGGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3085
Db 7135 GCGCCGATGATCATGAGACGCGACCGTTCGCGAGATGAGCGGTTCATCTATCTATCTG 7076
Qy 3086 CTGCGCTCAGCGCGATGCGCTATTGATTTGAAGATACCATTAAGTTAACGACCGCGG 3145
Db 7075 CTGCGCTTCACTGCTCCACCGCATCTGATCGAGGATACGCGCTACAGCGCGGCGAT 7016
Qy 3146 CTGGCGGAGAACACCGCTCGTCAGCACATCGCGCTATGCGCAATTCAGCAAGGCTGAG 3205
Db 7015 CTGGACGATGCGCGCTGGCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6956
Qy 3206 CTGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3265
Db 6955 GGGCAGGAGATGCGCGCGGCAAGGGGCTATCTGCGCATCGCGCTGCGCTGCGCTG 6896
Qy 3266 GATCGATTCTGCAACAGCAGCGCGCAAGGTGCGAGCGGCTGCGCGCGGCTGCTG 3325
Db 6895 GGTCTTCTGCGCGCACGCGCGAGGGGCGGCTGCGCGTGGGCTGGGGCAGGGCTGTT 6836
Qy 3326 CATGCAACACCGGTTACTCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3385
Db 6835 CACCCGCTCAGCGGATATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6776
Qy 3386 CTGTTGCGCACCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3445
Db 6775 CCGGACCTGACGACGCGCTGCGCGCTGCGCGGCTGCGCGCTGCGCGCTGCGCG 6716
Qy 3446 TGGCGGCAACAGCGATTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3505
Db 6715 GATCGCG---ACCGCTTCTGCGGCTGCTGAAACCGGATGCTGTCGCGGCTGCCG 6659
Qy 3506 CAGCAGCGCTGCGCGCTGATGCAACGTTTTTACCGGCTCGATGCGCGGTTAATTA 3565
Db 6658 GACCGTCTGCTATCGCTGCTGCGCGGTTCTACCGCTGCGCGACCGCTGATCGAGCG 6599
Qy 3566 TTTTACCGCGGCAATGCGCGCTGCGCGATATAAACCAGGATTTCTGTCGCGGAGCG 3625
Db 6598 TTCTATGCGCGGCTGACATTTGCGCGACCGGCTTCCGATGCTCACCAGCGCGCG 6539
Qy 3626 GTGCCCATCGGTGAAGCGCTGCGCGCTGTTGAATTCGTGCAACACGAGGAGAAAA 3685
Db 6538 ATTCCGCTGTCGACGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6479
Qy 3686 TGAAACGCACTTATGTTGCGCAGGCTTTTGGCGGCTGCGGCTGCGGCTGCGGCTG 3745
Db 6478 TGAGTTCCGCACTGCTGCTGCGCGGCTTTCGCGGCTTGGCTTGCATCCGCTG 6419
Qy 3746 AAGCGCGGCGATACCAACACTTCTGAGACGCGCGCAAAACGCGCGGAGCGCGCT 3805
Db 6418 AATCGCGGCGATTCGCGACCACTGTCGAGCGCGCGCAAGCGCGCGCGCGCT 6359
Qy 3806 ATGTTGAGGACAGTGGCTTTACTTTCGATGCGCGGACCGCGCTGATCACCGATCCA 3865

Db 6358 ATGCTGGAACGATCAGGCGCACGCTCTTCGATGACGCGCCGACGCGTCTGTGACCGACCCCG 6299
 Qy 3866 GGGCCATCGAAGAGTTGTTACGCTGGCAGGAAATCGCTCAGCGATTACGTCGAGCTGA 3925
 Db 6298 ACAGCCTGCGAGAGTGTGGGCCCTCAGCGGCAACCGATGAGGCGTGAAGCTGCG 6239
 Qy 3926 TCCCGGTAAAGCCCTTCTATCGCCTGTGCTGGGAAGATGGAAACAGCTTGAATACAGCA 3985
 Db 6238 TCCCGGTCTCGCCCTTCTACCGCTGACATGGGCGGACGCGCGAGCTTCGAATACCTGA 6179
 Qy 3986 ATAAATCAGCCGCTGTGAGCAGAGATGCGCAGCTCAATCCGCAAGATGTAGAAGCT 4045
 Db 6178 ACAGCAGCAGCAGCTGATCCGCCAGGTGCGCTCTCAATCCGCCGATGTGATGCT 6119
 Qy 4046 ATCGTCAATTTCTTGCTATTCACTGTAAGTATTTAGAGAGGGTTATCTGAAACTCGGCA 4105
 Db 6118 ATCGCGCTTCCAGNATACGCGGAGGTCTATCGCGAGGGTATCTGAAGCTGGGGA 6059
 Qy 4106 CGGTGCGGTTTCTGAGGTGCGTGAATGCTCGGCTGCGCGCGCAGTTGGAGCTGCG 4165
 Db 6058 CCACGCCCTTCTGAAAGCTGGGCGAGATGCTGAACGCGCGCGCGCTGATGCGCCTGC 5999
 Qy 4166 AAGCATGGCAGCGCTACAGCATGTCGCGAAATTTATTCAGACGATCATCTGCGTC 4225
 Db 5998 AGGCATACCGCTCGGTCCACAGATGTTGGCGCTTCAATCAGACCGCATCTGCGGC 5939
 Qy 4226 AGCGGTTTCTTCCACTCATTTGCTGTTGGGGGTAATCCTTTTGAAGCTCATGATCT 4285
 Db 5938 AGGCTTCTGTTCCACACGCTGCTGTTGGGGAAACCGCTTTTTCAGCACGCTGATCT 5879
 Qy 4286 ATACCTTAATTCATCGCTGAGCGTGAATGGGGCGTGTGTTCCGCGCGCGCACCG 4345
 Db 5878 ATGCGCTGATCCATGCGCTGGAACGCGCGCGCGCGCTCTGTTGCGCAAGCTGCTGA 5819
 Qy 4346 GCGCGCTGTGCAAGGCGATGCGCGCACTGTTTCAGAGCATTTGGCGCGGAGCTGTTACTGA 4405
 Db 5818 ACCAGCTGGTTCGCGGCGATGCTGCGCTGTTGAGCGCTTTCGCGGCAAGCTGCTGA 5759
 Qy 4406 ATGCGGAAGTGAGCAGCTGGAACAGCGGCAATCGCATTCAGCGGCTTCAGTTAGAG 4465
 Db 5758 ATGCGCGGCTCAGCGGATCGACACCGAGGCGATCGCGCACGGGCGTCACTGCTGG 5699
 Qy 4466 GCGGACGACGCTTGCATGCGCGCTGTCCTCAATGCGGAGCTGTTGTCATACCTAGC 4525
 Db 5698 ACGGGCGGAGTTGCGCGGATGAGTGGCGGAGTGGCGGAGTGGCGGATGCAAGCTATC 5639
 Qy 4526 ACAAACTGCTTTCGCCACCATCCGCTGGCAATGAAGCGTGGACATCGCTGAAGCTAAGC 4585
 Db 5638 GCGACCTGCTGGGCGATACCGCGCGCGCGCGCACCAAGCGCGGATCCTGAACCGGAGC 5579
 Qy 4586 GCATGAGCACTCGCTGTTGCTACTTATTTGGCCTGAAATCAGCGCGCATGAACAGCTCG 4645
 Db 5578 GCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5519
 Qy 4646 GCGACCAACGCTTGTGTTGGCGCGGTTATCGTGAGTTGATCGATGAGATTTTCAACA 4705
 Db 5518 CCCACCAAGCGTCACTTCGCGCGCGCTCAAGAGGGCTGTTGAACGAGATCTTCAACG 5459
 Qy 4706 GAGCAGCTGCGCAGCATTTTCACTTACCTGCAAGCGCGCTGCGAGCGCATCCGT 4765
 Db 5458 GGGCAGCGCTGCGGACGATTTCTGATGATCTGATTCGCTGCGCTGCGGATCCCA 5399
 Qy 4766 CGCTGCGACCGCGCTGCGGAGCTTTTATGTTGTTAGCGCGCGGTCGCGATCTGGCA 4825
 Db 5398 GCTTGGCCCCGAGGGATGTCACGATAGCTTCTTGGCGCGCTTCCGCTCTGGGCGC 5339
 Qy 4826 CGCTGACATCGACTGGCAACAGGAAGACCGCGCTTGGCGGATCGGATTTTCTTATC 4885
 Db 5338 GCGCGGATGCTGATTTGGGAAGCGGAGCGCGGCTATGCGGAGCGCATCTTCGAGGAC 5279
 Qy 4886 TGGAGCAGCATACATCGCGGATTAAGTCAAGATTAAGTGAACACAGAAATGTTTACGC 4945
 Db 5278 TGGAGCGCGCGGCGCATCCCGACCTCGGCAAGCACTGACCGCTGACCGCGCATCTTCAGCC 5219

Qy 4946 CGTTTATTTTTCGCGACACGCTGTCATGCCATCAGGCTCGCGCTTTTCGCTGGAGCCGA 5005
 Db 5218 CGCCGATTTTACGACACCGAACCTGTGCGGCCATCAGCGAGCGCTTCTCGTTCGAGCCGA 5159
 Qy 5006 TTTTGAACGAAAGCGCTGTTTCCGCGCGCATAAACCGGATGCGGATATCAGCAATCTCT 5065
 Db 5158 TCTGACGCAATCCGCTGTTTCCGCGCGCATAAACCGGACCGCGCATCCGAACTTCT 5099
 Qy 5066 ATCTGTTGGTGGCTGAGCATCAGGCGCGGCGTGCCTGCGCTGATCGGTTTCGGGCA 5125
 Db 5098 ACATCTGTTGGGCGGCGACGCTGCGGTTGCGGCTATCCCGGTTGCTGTTGCGAGCGCA 5039
 Qy 5126 AGGCCACCGCAGGCTGATGCTGGAGATCGCGCGAATGAATCGACACCTTTACTTGA 5185
 Db 5038 AGGCCACCGCGCAGGTCATGCTGCGAATGCGGCTGCGCATGCGATCTGACGCGAC 4979
 Qy 5186 GCAAGTAAACGAAACCATGCGGCTGGCTGAGAGTTTCGCCACCGCGCGCAAGCTGT 5245
 Db 4978 TTTCCGAAGCG---GCCATCGCGAGGTTTCGAAAGCTTCGCGCAGCGCGCGCTGAT 4922
 Qy 5246 TGATGCAACGACGCGCGCAGCAGCTGATGCTGATGCTGATGCTGCTGCTGCTGCTGCTG 5305
 Db 4921 GCGCGCGGCTGCGCGAGGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4862
 Qy 5306 TGTGATGATGGGCAACGCTGGCGGAAAGCGGCGACGAGCATGCGCTCGAAGACGCGCA 5365
 Db 4861 CGTGATCGACGCGGAGGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4808
 Qy 5366 GGCAGTATGAGCATCTGCAAAATTTGAAACCGCGCGCTACAGCGCGCGCGCATGGA 5425
 Db 4807 GCGCGGCTGGGCGCTGCGCGGACACGCTGCGCGCTGCTGCGAGGACGCGCGCT 4748
 Qy 5426 TGAACCGCGCTTTAGGCGCTTTCAGGAAGTGGGATCATTCACAGCTGCGCAACACT 5485
 Db 4747 GTCGCGCGCTTTCGCGCGCTGCGCAGCTGCGCGCGCGCATGATTTCCCGACCTTTG 4688
 Qy 5486 GCGCTTTCATCTGNAAGCTTTCGCTATGATGATGACGCAACGAACTTACGCGAGCTT 5545
 Db 4687 GCGGATGACCTGATGAGAGGTTTCGCGATGGAATGTCGCGATGCGGATGCGGATACGCGCT 4628
 Qy 5546 CGATGACAGCTGCTTACTGCTATCAGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5605
 Db 4627 GATGACGCTGCTGNAATTTCTACCACTGCGCGGCTGCTGCGGCTGATGATGCGCG 4568
 Qy 5606 CTTAATGGGCTGCGCGGAGGAGCGGCTGCTGATCAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTG 5665
 Db 4567 GGTGATGCGCTGCGAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4508
 Qy 5666 CCAGCTCACTAACATTTGCGCGCATTTGATAGAGATGCGGAAATGCTGCTGCTGCTGCTGCTGCTG 5725
 Db 4507 CCAGCTGAGCAACATCTGCTGCGGAGCTGATCGAGATGCGGATGCGGCTGCTGCTGCTGCTGCTG 4448
 Qy 5726 GCGCAATCTTGGCTGATCAGCGGAGTTAGCGCGATAGCTGATGCTGCTGCTGCTGCTGCTGCTG 5785
 Db 4447 GCTTGGCGACTGGCTGGCGAGGCGG-----GGCGAGCTGAGGCTGCTGCTGCTGCTGCTGCTG 4397
 Qy 5786 TGTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5845
 Db 4396 TTTGAGCGCTGCTTATTTCCGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4337
 Qy 5846 CTTGCGCGCATCCGCTTTTACCGGCTTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5905
 Db 4336 CTTGCGCGCGGAGGCTTTCGCGCATCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4277
 Qy 5906 TCGCGGCTTTATTCGCGAAATTTGGCTCAAGTTGAGCAGCAGCGGCTGCTGCTGCTGCTGCTGCTG 5965
 Db 4276 GCTGCTGCTATCTATTCGCGCAATCGGAGCGCGATTCGCGAGGCTGCGCGCGCTGCTGCTGCTG 4217
 Qy 5966 TTCACGCGAGCGCACAGTAAAGGTGAAAGTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6025
 Db 4216 CCAGCGGATCAGCACGCTCGNAGGCTGCCAAGATCGGCTTCTGCGCGCGGAGGCTTGA 4157

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ladas & Parry
STREET: 26 West 61 Street
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10023
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720Kb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: N/A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/783,705A
FILING DATE: 19911023
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1-103078
FILING DATE: 21-APR-1989
APPLICATION NUMBER: JP 2-53225
FILING DATE: 05-MAR-1990
APPLICATION NUMBER: US 07/519,011
FILING DATE: 19-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Schwadron, Janet I.
REGISTRATION NUMBER: 33,778
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-708-1935
TELEFAX: 212-246-5959
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1479 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid (plasmid DNA)
US-07-783-705A-10

Query Match 11.8%; Score 826.2; DB 1; Length 1479;
Best Local Similarity 72.6%; Pred. No. 2.5e-216;
Matches 1068; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

QY 3685 ATGAACGCACTATGTGATGGCGCAGGCTTTGGCGGCTGGCGCTGGCGATTCGCTG 3744
DB 1 ATGAACCACTACGCGTAATGGTGCAGGCTTCGGTGGCTGGCACTGGCAATTCGTCTA 60
QY 3745 CAAGCGGGGCGATACCAACCACTTACTCGAGCAGCGCAGCAAAACCGGCGGAGCGCC 3804
DB 61 CAAGCTGCGGGGATCCCGCTTACTGCTTGAAACACGTGATAAACCCGGCGGTGGGCT 120
QY 3805 TATGTGTTTGGAGCAGTGGCTTTTACCTTCGATGCGGACCCACCGTGATCACCGATCCC 3864
DB 121 TATGCTACAGGATCAGGGGTTTACCTTTGATGAGGCGCCGAGGTTATCACCGATCCC 180
QY 3865 AGCGCATCGAAGAGTTGTCAGCTGGCAGGAAATGCTCAGCGAATTCGTCGAGCTG 3924
DB 181 AGTGCCATTGAAGAACTGTTTGCACTGGCAGGAGAAACAGTAAAGAGTATGTCGAACTG 240
QY 3925 ATGCGGTAAACCGCTTCTATCGCTGCTGCGAGATGGCAACAGCTGATACGAC 3984
DB 241 CTGCGGGTATACGCGGTTTACCGCTGTGTGGAGTCAGGGAAGGCTCTTAAATACGAT 300
QY 3985 AATAATCAGCGCTGCTGGAGCAGCAGATCGCCACCTTCAATCCGCAAGATGTAGAAGC 4044
DB 301 AAGCATCAACCGGCTCGAAGCGCAGATTCAGCAGTTTAAATCCCGCGATGTCGAAGT 360
QY 4045 TATCGTCAATTTCTTGCCTATTACGCTGAAGTATTATAGAGAGGTTATCTGAAATCTGGC 4104
DB 361 TATCGTCAGTTTCTGACTATTACCGCGGGTGTGTTAAAGAAAGGCTTATCTAAAGCTCGT 420
QY 4105 ACGTGCCGCTTCTGACAGTGGTGACATGCTCGCGTCCGCGGAGTTGGAGCTCTG 4164

RESULT 7
US-08-095-726-7

DB 421 ACTGTCCCTTTTATCGTTTACAGACATGCTTCGGCGCGACCTCACTCACTGGCGAACTG 480
QY 4165 CAAGCATGCGCAGCGTCTTACAGCATGTTGGCGAAATTTATTACGACGATCATCTGCGT 4224
DB 481 CAGGCATGAGAAAGCGTTTACAGTAAGGTTGCGAGTTACATCGAAGATGAACATCTGCGC 540
QY 4225 CAGGCGTTTCTTCCACTCATTTGCTGGTGGGCGGTAACTCTTTGCAAGCTCATGATC 4284
DB 541 CAGGCGTTTCTTCCACTCGCTGTTGGTGGCGGCAATCCCTTCGCCACCTCATCCATT 600
QY 4285 TATACCTTAATTCATGCGCTGAGCGTGAATGGGGCTGTGGTTCCGCGCGCGCGCAC 4344
DB 601 TATAGTTGATACACGCGCTGAGCGTGAAGTGGGGGCTGTGGTTCCGCGGTGGCGCAC 660
QY 4345 GCGCGCTGCTCAGGGCATGCGGCATGTTTCGAGGACTTGGGCGCGAGCTGTCTACTG 4404
DB 661 GCGCATTAGTTTACGGGATGATAAAGCTGTTTCAGGATCTGGGTGCGAAGTCGTGTTA 720
QY 4405 AATGCGAAGTACGCGCTGAGAAACGAGCGGCAATCGCATTTAGCGGCTTCAGTTAGAG 4464
DB 721 AAGCGCAGAGTCAGCCATATGGAACGACAGGAAACAAGATTGAAGCCGTGCAATTAGAG 780
QY 4465 GCGGACGACGCTTCGATGCGCGCTGTGGCTTCCAATGCCGACGCTGTGCAATACCTAC 4524
DB 781 GACGCTCGCAGGTTCTGACGCAAGCCGTCGCGTCAATGCAAGTGTGTTTATACCTAT 840
QY 4525 GACAAATGCTTTCGCCACCATCCGCTGGCAATGAACGTCGACATFCGCTGAAGCGTAAG 4584
DB 841 CGCGACCTGTTAAGCCAGCACCTTCGCGGGTTAAGCAGTCCAAACAAATGCGAGACTAAG 900
QY 4585 CGCATGAGCACTCGCTGTTTGTACTTATTTTGGCTGAATCAGCGCGATGAACAGCTC 4644
DB 901 CGCATGAGTAACCTCTGTTTGTGCTTATTTTGGTTTGAATCACCATCATGATCAGCTC 960
QY 4645 GCGCACCAACGCTGTTTGGCGCGCTTATCGTGAGTGTGATCGATGAGATTTTCAAC 4704
DB 961 GCGCATCACACGCTTGTTCGCGCGCTTACCGGAGCTGATGACGAAATTTTAT 1020
QY 4705 AGCAGCAGCTGCGCAGAGATTTTCACTTACCTGACGCGCCCTGACGAGCAGATCCG 4764
DB 1021 CATGATGGCTCGCAGAGACTTCTCATTATCTGACGCGCCCTGTGTACCGGATTCG 1080
QY 4765 TCGCTGGCACCGCGCTGCGGAGCTTTTATGTTTGTAGCGCGCTGGGATCTCGGC 4824
DB 1081 TCACTGGCGCTTGAAGGTTGCGGCAAGTTACTATGTTGTCGCGCGGTGCGCAATTTAGGC 1140
QY 4825 ACCGCTGACATCGACTGGCAACAGGAAGGACCGCGCTTCGCGCATCGAATTTTGTCTAT 4884
DB 1141 ACCGGAACCTGACTGACGCTTACGAGTTCAGGAGTTCAGCGCATGATTTTGGGTAC 1200
QY 4885 CTGGAGCAGCTATACATGCGGGATTTACGTACAGCAATTTAGTACACACAGAAATGTTTACG 4944
DB 1201 CTGGAGCAGCTATACATGCTGGCTTACGAGTTCAGGAGTTCAGCGCATGATTTTACG 1260
QY 4945 CGGTTGATTTTGGCAGCAGCTGATGCCATCAGGCTCGCGCTTTCGCTGGAGCGC 5004
DB 1261 CGGTTGATTTTGGCAGCAGCTTAAATGCTTATCATGCTCAGCTTTCGCTGGAGCGC 1320
QY 5005 ATTTTGAACCAAGCGCTGTTTCGCGCGCATAAACCGGATGCGCATTCAGCAATCTC 5064
DB 1321 GTTCTTACAGAGCGCTGTTTGGCGCATTAACCGGATTAACCAATTAATATCTC 1380
QY 5065 TATCTGTGGTGGTCCGGTACGATCCAGCGCGGCGTCCCGGCGTGTGATCGGCTCGGC 5124
DB 1381 TACCTGGTGGCGAGGACGATCCCGCGCAGGCAATTCCTGGCGCTCATCGGCTCGGCA 1440
QY 5125 AAGGCCACCGCAGGCTGATGCTGGAGATC 5155
DB 1441 AAGGCCACAGCAGGTTTGTGCTGGAGATC 1471


```
Sequence 7, Application US/08095726
Patent No. 5530188
GENERAL INFORMATION:
APPLICANT: Ausich, Rodney L
APPLICANT: Brinkhaus, Friedhelm L
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H
APPLICANT: Yarger, James G
APPLICANT: Yen, Huel-Che B
TITLE OF INVENTION: Beta-Carotene Biosynthesis in
Genetically Engineered Hosts
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amoco Corp., Patents and Licensing Dept
STREET: 200 E Randolph St
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60680-0703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,726
FILING DATE: 21-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/785,566
FILING DATE: 30-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Galloway, No. 5530188val B
TELEPHONE: 3128567180
TELEFAX: 3128564972
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1518 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-095-726-7

Query Match 10.28; Score 713.4; DB 1; Length 1518;
Best Local Similarity 67.78; Pred. No. 2.4e-185;
Matches 1034; Conservative 0; Mismatches 481; Indels 12; Gaps 2;

Qy 3680 AAAAAATGAAACGCACTATATGATTGGCGCAGGCTTTGGCGCCTGGCGCTGGCGATTC 3739
Db 2 AAACCATGGAATAAACCGTTGTGATTGGCGCAGGCTTTGGTGCGCTGGCGCTGGCGATTC 61
Qy 3740 GCCTGGAAGCGGGGATACCAACACCTTACTCGAGCAGCGCGCAACACCGGGCGGAC 3799
Db 62 GCCTGCGAGGCGGAGGATCCCAACCTGCTGCTGGAGCGGCGGCAAGACCGCGGGTC 121
Qy 3800 GCGCCTATGTTTGGAGCAGTGGCTTTTACCTTCGATGCGGACCGACCGAGTCAACCG 3859
Db 122 GGGCCTACGCTGCGCAGTACAGGCGCTTTACCTTTGACGCGCGGCGGCGAGCTGATCACCG 181
Qy 3860 ATCCAGCGGCATCGAAGAGTGTTCACGCTGGCAGGAAATCGCTCAGCGATTAAGTCG 3919
Db 182 ATCCAGCGGCTTGGAGCGCTGTTTCACTGCGCGGCGGCGGCGGATGAGGATTAAGTCA 241
Qy 3920 AGCTGATCGGTAACGCCCTTCTATCGCTGCTGCGAAGATGGCAACAGCTTGATT 3979
Db 242 GGCTGCTCGCGGTAACCCCTTCTACCGACTCTCTGGAGTCCGGAGAACCTCGACT 301
Qy 3980 ACGACAATAATCAGCGCGTCTGGAGCAGCAGATCGCAAGTTCATTCGCAAGATGATG 4039
Db 302 ATGCTAACGACAGCTTCGAGCTTGAGGCGCAGATTACCCAGTTCAACCCCGCGACGTCG 361
```

Db 1433 TCGCGAAGACCGCCAGCCCTGATGATTGAGATC---TGCAATGAGCCAAACCGCCGCT 1489
Qy 5180 ACTTGAGCAAGTAACGCAAAACCATGGC 5206
Db 1490 GCTTGACACGCGCACGACCATGCG 1516

RESULT 8
US-08-096-043-7
; Sequence 7, Application US/08096043
; Patent No. 5530189
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Hueli-Che B
; TITLE OF INVENTION: Lycopene Biosynthesis in
; TITLE OF INVENTION: Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,043
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,568
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530189val B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-096-043-7

Query Match 10.2%; Score 713.4; DB 1; Length 1518;
Best Local Similarity 67.7%; Pred. No. 2.4e-185;
Matches 1034; Conservative 0; Mismatches 481; Indels 12; Gaps 2;

Qy 3680 AAAAATGAACGCACCTATGATTTGGCGCAGGCTTTGGCGCCTTGGCGCTGGCGGATTC 3739
Db 2 AAACCATGGAATAAACCGTTGTGATTGGCGCAGGCTTTGGTGGCCTGGCGCTGGCGGATTC 61

Qy 3740 GCTCTGAAGCGCGCGGATACCAACACCTTACTCGAGCAGCGCGACAAACCGCGCGGAC 3799
Db 62 GCTCTGAAGCGCGGAGGATCCCAACCGTACTGCTGAGCAGCGGGAACGCCCGCGGTC 121

Qy 3800 GCGCCTATGTTTGGAGCAGTGGCTTTACCTTCGATGCGGACCCACCGGATCACCG 3859
Db 122 GGGCCTACGCTCGCATGACCAAGGCTTTACCTTTGACGCGCGGCGAGCGGTGATCAACG 181

Qy 3860 ATCCGAGCGCATCGAAGAGTTGTTACGCTGGCAGGAAATCGCTCAGGGATTACGTG 3919
|||||

Db 182 ATCTACCCGCGCTTGAGGCGCTGTTCAACCTGCGCGGAGGCGCATGGAGGATTACGTCA 241
Qy 3920 AGCTGATGCGGTAAGCGCCCTTCTATCGCTGTGCTGGGAAGATGGCAACAGCTTGATTC 3979
|||||
Db 242 GGTGCTGCGCGTAAACCCCTTCTACCGACTCTGCTGGGAGTCCGGGAAAGACCTCGACT 301
|||||
Qy 3980 ACGACAATATCAGCGCGTCTGGAGCAGCAGATCGCCAGCTTCAATCGCAAGATGTAG 4039
|||||
Db 302 ATGCTAAACGACAGCTTCGAGCTTGAGGCGCAGATTACCCAGTTCAACCCCGCGAGCGTCG 361
|||||
Qy 4040 AAGGCTATCGTCAATTTCTTCTGCTTATTCAGCTGAAGTATTTAGAGAGGGTATCTCAAAAC 4099
|||||
Db 362 AGGGCTACCGGCGCTTCTGCGCTTACTCCAGCGGCTATTCAGAGAGGATATTTGGGCC 421
|||||
Qy 4100 TCGGCACGCTGCGGTTTCTGCAAGTGCCTGATGATGCTGCGGCTGCGCGCAGTGGGAC 4159
|||||
Db 422 TCGGCAGCGTGCCTTCTCTCTTTTCGCGACATGCTGCGCGCGCGCGCGCAGCTGCTTA 481
|||||
Qy 4160 GTCTGCAAGCATGGCGCAGCGCTTACAGCATGGTGGCGGAAATTTATTTCAGGAGCATC 4219
|||||
Db 482 AGCTCCAGCGCTGGCAGAGCGCTTACCAGTTCGGCTTTTCGCGCTTATTGAGGATGAGCATC 541
|||||
Qy 4220 TCGCTCAGCGCTTTCTTCTCACTCATTTGCTGGGCGGTAACTCTCTTTTGCACCGTCAT 4279
|||||
Db 542 TCGCGAGCGCTTCTGTTTCTCACTCCCTGCTGTGAGCGCGCAACCCCTTACCACCTCGT 601
|||||
Qy 4280 CGATCTATACCTTAATTCATGCGCTGGAGCGTGAATGGGCGCTGTGTTTTCGCGCGGCG 4339
|||||
Db 602 CCATCTACACCTGATCCAGCGCTTGAGCGGAGTGGGGGCTCTGTTTCCCTGAGGGCG 661
|||||
Qy 4340 GCACGCGCGCTGCTGAGGCGATGGCGGACTGTTTCAGGACTTGGGGGGCGAGCTGT 4399
|||||
Db 662 GCACCGGCGCTGCTGTAACGCGCATGGTGAAGCTGTTTACCAGTCTGGGCGGGAGATCG 721
|||||
Qy 4400 TACTGAATGCCAAGTGAAGCGCAGCTGGAACCCAGCGCAATCGCATTTAGCGGCTTCAGT 4459
|||||
Db 722 AACTCAACCGCGGCTCGAAGAGCTGTGCTGCGCCGATTAACCGCTGAAGCAGCTCGGC 781
|||||
Qy 4460 TAGAGGGCGGACGAGCTTTCGATGCGCGCGCTGTGGCTTCCAATGCCGAGCTGGTGAT 4519
|||||
Db 782 TCGCGGATGCTCGGATCTTTGACACCGACGCGCTGAGCTCGAAGCTGAGTGAACA 841
|||||
Qy 4520 CCTACGACAACTGCTTCCGACCATCCGCTGGCAATGAACGCTGGACATCGCTGAGC 4579
|||||
Db 842 CCTATAAAAGCTGCTCGCGCACCATACCGGTGGGCGAGAACGCGGCGGCGCGCTGGAGC 901
|||||
Qy 4580 GTAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCCCTGAATCAGCCCATGAAC 4639
|||||
Db 902 GCAAGCATGAGCAACTCGCTGTTTGTGCTCTACTTCGCGCTGAACACGCTCATTTCCC 961
|||||
Qy 4640 AGCTCGCGCACCAACCGCTCTGTTTGGCCCGGTTTATCGTGAAGTTGATGATGAGATTT 4699
|||||
Db 962 AGCTGCGCACCATACCATCTGTTTGGTCCCGCTACCGGAGCTGATCGACGAGATCT 1021
|||||
Qy 4700 TCAACGAGCAGCAGCTGGCAGAGATTTTCACTTTACCTGCAAGCGGCTGCGAGCG 4759
|||||
Db 1022 TTACCGGACGCGCTGGCGGATGACTTCTCGCTCTACCTGCACTCGCCCTCGGTGACCG 1081
|||||
Qy 4760 ATCCGCTGCTGCGCCCGCGCTGGCGAGCTTTTATGTTGTAGCGCGGTCCCGCATC 4819
|||||
Db 1082 ATCCCTCGCTCGCGCTCCCGCTGCGCAGCTTCTAAGTGTGCGCCCGGTGCGGCATC 1141
|||||
Qy 4820 TCGGCACCGCTGACATCGACTGGCAACAGGAAGGACCGCGCTTTCGCGATTCGAATTTTG 4879
|||||
Db 1142 TTGGCAACGCGCTGGACTGGCGCAGAGGGCGGAAGCTGCGGACCGCATCTTTG 1201
|||||
Qy 4880 CTTATCTGAGCAGCATACATGCGCGGATTTAGTCAGCAATTTAGTGCACACAGATGT 4939
|||||
Db 1202 ACTACCTTGAAGAGCGCTATATGCGCGCTGCTGAGCAGCTGGTGAACCCAGCGATCT 1261
|||||
Qy 4940 TTACGCGCTTGTATTTTCGCGACGCTCATGCGCCCATCAGCGCTGGGCTTTTCGCTGG 4999
|||||
Db 1262 TTACCGCGCAGACTTCACGACAGCTTGGATCGCATCTTTGGGATCGCTTTTTCATCGAGC 1321
|||||

QY 5000 AGCGATTTTACGCAAGCGCCGTGGTTCCGCGCGCATAAACCGCGATGCCGATATCAGCA 5059
Db 1322 CGCCTTCGTTGACCCCAAGCTTGTTCGCCGAACCGCACGACATTC-----AA 1372
QY 5060 ATCTCTATCTGGTGGGTGCCGGTACGATCCAGCGCGGGGGTGCCTGGCGGTGATTCGGTT 5119
Db 1373 ACCTCTACCTGGTGGCGCAGGTACTCACCTGGCGGGGCATTCCTGGCGTAGTGGGCC 1432
QY 5120 CGGCAAGGCCACCGCCAGGCTGATCTGGAGGATCGCGCGAATGAATCGACAGCCTTT 5179
Db 1433 TCGCGCAAGCACCGCCAGCCTGATGATGAGGATC---TGCATAGCCAAACCGCGCT 1489
QY 5180 ACTTGAGCAAGTAACGCAAAACCATGGC 5206
Db 1490 GCTTGACCAAGCCAGCAGACCATGGC 1516

RESULT 9
US-08-096-623A-7
; Sequence 7, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Host
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Welsh & Katz, Ltd.
; STREET: 120 S. Riverside Plaza, 22nd Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,623A
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,061
; FILING DATE: 09-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/662,921
; FILING DATE: 28-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/562,674
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/525,551
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,613
; FILING DATE: 02-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: AMO-006.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 655-1500
; TELEFAX: (312) 655-1501
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1518 base pairs

QY 4580 GTAAGCGCATGAGCACTCGCTGTTTGTACTCTATTTTGGCCTGATCAGCGCGATGAC 4639
Db 902 GCAAGCATGAGCACTCGCTGTTTGTCTTACTTTCGGCTGAAACGAGCTCATTTCC 961
QY 4640 AGCTCGCGACACACCGCTGTTTGGCCGCTTATCGTGTGATGATGATGATTT 4699
Db 962 AGCTGGCGACCATACCATCTGTTTGGTCCCGCTACCGGAGCTGATCGAGATCT 1021
QY 4700 TCAACGAGCAGCAGTGGGAGACGATTTTCTACTTCTGACCGCGCTCGACGAGCG 4759
Db 1022 TTACCGCGCAGCGCTGGCGGATGACTTCTCGCTCTACCTGACCTCGCCCTCGTGACG 1081
QY 4760 ATCCGCTGCTGACCGCGCTGCGGAGCTTTTATGTTAGCGCGCTGCGGATC 4819
Db 1082 ATCCCTCGCTGCGCGCTCCCGCTGCGGAGCTTCTACGTGCTGCGCGCTGCGGATC 1141
QY 4820 TCGGACCGCTGACATGCACTGCAACAGAGAGCGCGCTTGGCGGATCGAATTTTG 4879
Db 1142 TTGGCAGCGCGCTGACTGGCGCAGAGGGCGGAGCTGCGGACCGCATTTTG 1201
QY 4880 CTTATCTGAGGAGCACTACATGCGGGGATTAAGTCAGCAATTAAGTCAGCAAGATG 4939
Db 1202 ACTACCTTGAAGAGCGTATATGCGCGCTGCTGAGCCAGCTGCTGACCCAGCGATCT 1261
QY 4940 TTACGCGCTGATTTTTCGCGACAGCTGATGCCCATCAGCGCTCGGCTTTTCTGCTGG 4999
Db 1262 TTACCGCGGAGACTTCAACACAGCTTGGATCGGATCTTGGGATCGCTTTTTCATCGAGC 1321
QY 5000 AGCCGATTTTGAACGAGCGCTGTTCCGCGCATTAACCGCGATGCGGATATCAGCA 5059
Db 1322 CGCTTCTGTTACCCAGAGCTTGTTCGCGCAACCGGACAGCATTC-----AA 1372
QY 5060 ATCTCTATCTGTTGGTCCGCTGAGCATCCAGCGCGCGGCTGCGCGCTGATCGGTT 5119
Db 1373 ACTCTACTGTTGGCGCAGGTACTCACTTGGCGCGGCTTCTGCGCTAGTGGGCG 1432
QY 5120 CGGCCAGCGCACCGCGCTGATGCTGAGGATCGCGCGGATGCAATGCAAGCTTT 5179
Db 1433 TCGCGAAGACCGCGCTGATGATGAGGATC---TGCAATGAGCAACCGCGCT 1489
QY 5180 ACTTGAGCAAGTAACCAACCATGCG 5206
Db 1490 GTTGGACAGCGCACGACCATGCG 1516

RESULT 10

US-08-095-726-9

Sequence 9, Application US/08095726

Patent No. 5530188

GENERAL INFORMATION:

APPLICANT: Ausich, Rodney L
APPLICANT: Brinkhaus, Friedhelm L
APPLICANT: Mukharji, Indrani
APPLICANT: Proffitt, John H
APPLICANT: Yarger, James G
APPLICANT: Yen, Huel-Che B

TITLE OF INVENTION: Beta-Carotene Biosynthesis in

OPERATING SYSTEM: Genetically Engineered Hosts

NUMBER OF SEQUENCES: 79

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amoco Corp., Patents and Licensing Dept

STREET: 200 E Randolph St

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60680-0703

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/095,726

FILING DATE: 21-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICANT: US 07/785,566
FILING DATE: 30-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Galloway, No. 5530188val B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3128567180
TELEFAX: 3128567180
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1522 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-095-726-9

Query Match 10.1%; Score 708.4; DB 1; Length 1522;

Best Local Similarity 67.6%; Pred. No. 5.6e-184;

Matches 1029; Conservative 0; Mismatches 481; Indels 12; Gaps 2;

QY 3685 ATGAACGACCTATGATTTGGCGCAGCTTTGGCGCTCGCTGCGATTCGCCCTG 3744
Db 11 ATGGAATAAACCGTTGTGATTTGGCGCAGCTTTGGCGCTCGCTGCGATTCGCCCTG 70
QY 3745 CAAGCGGGCGGCATACCAACCACTTACTTCGAGCAGCGCGCAAAACCGGGCGGACGCGC 3804
Db 71 CAGGCGGCGAGGATCCCAACCGTACTGCTGGAGCAGCGGACAGCGCGGCGGCGC 130
QY 3805 TATGTTTGGAGCAGTGGCTTTTACCTTCGATGCGGACCAACCGGTGATCCCGATCCC 3864
Db 131 TACGCTCGCATGACAGCGCTTTTACCTTTGACGCGCGGCGGACGCTGATCACCAGTCT 190
QY 3865 AGCGCCATCGAGAGTGTTCAGCTGCGCAGGAGAAATCGCTCAGCGATTTACGTCGAGCTG 3924
Db 191 ACCTGCTTGGCGCTGTTTACCTTCGCGCGCAGCGCATGAGGATTTACGTCAGGCTG 250
QY 3925 ATGCGCGTAACCGCTTCTATCGCTGCTCGGAGATGCGCAACAGCTTGAATACGAC 3984
Db 251 CTGCGGTAAACCTTCTTACCGACTCTGCTGGAGTCCGGGAGACCTTCGACTATGCT 310
QY 3985 AATAATCAGCGCTGCTGGAGCAGCAGATCGCACTTCATTCGCAAGATGTAGAGGC 4044
Db 311 AACGACAGCTTCGAGCTTGGCGCAGATTAACAGCTTCAACCCCGCGCAGCTCGAGGCG 370
QY 4045 TATCGTCAATTTCTGCTTATTCAGTGAAGTATTTAGAGAGGTTATCTGAACTCGGC 4104
Db 371 TACCGCGCTTCTGCTTACTTCCAGCGGTATTTCCAGAGGGGATTTTGGCGCTCGGC 430
QY 4105 ACCTGCGCGTTTCTGAGGTGCTGACATGCTCGCGCTCGCGCGCAGTTGGGACGCTCG 4164
Db 431 AGCTGCGCTTCTCTCTTTTCGCGCATGCTGCGCGCGCGCGCGCTGCTTAAAGCTC 490
QY 4165 CAAGCATGCGCAGCTGTACAGCATGCTGGGAGAAATTTATTCAGAGCATCATCTGCGT 4224
Db 491 CAGGCGTGGCAGAGGCTTACCACTCGCTTTCCGCTTTTATTTAGGATGAGCATCTCGCG 550
QY 4225 CAGGCGTTTCTCTTCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4284
Db 551 CAGGCGCTTCTGCTTCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
QY 4285 TATACCTTAAATTCATGCGCTGGAGGTGAATGGGCGTGTGCTGCTGCTGCTGCTGCTGCT 4344
Db 611 TACACCTGATCCAGCGCTTGGCGGAGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670
QY 4345 GCGCGCTGCTGCGGATGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4404
Db 671 GCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730
QY 4405 AATGCGGAGTGAAGCGCTGCAACAGCGCATTCGATTTAGCGGCTTTCAGTTAGAG 4464

731 AACGCCGGTGAAGAGCTGGTGTGGCCGANTAAACCGGTAAGACCGAGTCCGGCTCGCG 790
4465 GCGCGACACGCTTCGATGCGCCCGCTGTGGCTTCAATGCGAGCTGGTGCATACCTAC 4524
791 GATGTGCGATCTTTGACACCGACCGCTGCTCGACGCTGAGCTGTGAACACCTAT 850
4525 GACAACTGCTTGGCCACATCCGCTGGCAATGAACCTGGCAATCCTCTGAAGGTAAG 4584
851 AAAAAGTCTCGGCACCATACCGTGGGCGAGAGCGGGCGGACGCTGGAGCGCAAG 910
4585 CCGATGAGCACTCGCTGTTGCTATCTATTTTGGCTGGAATCAAGCGCATGAACAGCTC 4644
911 AGCATGAGCACTCGCTGTTGCTATCTTGGCCCTGACCGGAGCTGATCGAGAGATCTTTACC 970
4645 GCGCACCAACACCTGTTTGGCCCGGTTATCGTGTGATCGATGAGATTTTCAAC 4704
971 GCGCACCAATACCATCTGTTTGGTCCCGCTACCGGGAGCTGATCGAGAGATCTTTACC 1030
4705 AGAGCAGCTGGAGAGATTTTCACTTTACCTGCGAGCGGCGCTGAGAGCGATCCG 4764
1031 GCGAGCGCTGGCGGATGACTTCTGCTCTACCTGCACTGCGCTGCGTGAACCGATCCC 1090
4765 TCGTGGCAGCGCGCGCTGCGGAGCTTTATGTTGTTAGCGCGGTGCGCATCTCGGC 4824
1091 TCGTGGCGCTTCCCGCTGCGGAGCTTCTAGTGTCTGCGCCCGCTGCGCATCTTGGC 1150
4825 ACCGTGACATCGATGCGCAACAGAGAGACCGCGCTTGGCGATCGAATTTTTCCTAT 4884
1151 AAGCGCGCTGAGCTGGCGGAGGAGGCGGCGAAGCTGCGGAGCGCATCTTTGACTAC 1210
4885 CTGGAGCAGCTACATCGCGGATTTACCTGCAATTAAGTACACACAGAGATTTTACG 4944
1211 CTTGAAGAGCGCTATATGCGCGCTGCGTGTGAGCAGCTGTGACCGAGATCTTTACC 1270
4945 CCGTTTGAATTTTGGCAGACGCTGATGCCCATACCGCTCGCGCTTTCGTGAGCGG 5004
1271 CGGAGACTTACGACACGCTTGGATCGGATCTTGGGATCGCTTTTATCGAGCGCT 1330
5005 ATTTGAGCAAGCGCTGTTCCCGCGCATACCGCGATGCGGATATCAGCAATCTC 5064
1331 TCGTTGACCAAGCTTGTTCGCGCAACAGCGGACACGATTC-----AAACCTC 1381
5065 TATCTGTTGGTCCCGTACGATCCAGCGCGCGGCTGCGCGCTGATCGGTTTGGCC 5124
1382 TACTGTGTCGCGAGTACTACCTTGGCGGCGATCTCTGCGTGTGCGCTCGCC 1441
5125 AAGGCCACCGCGCTGATGCTGGAGATGCGCGCGAATGAATCGACAGCTTTACTTG 5184
1442 GAAAGCACCGCGCTGATGATGAGGATC---TGCAATGAGCCAAACCGCGCTGCTTG 1498
5185 AGCAAGTACGCAACCATGCG 5206
1499 ACCAGCGCAGTCGACCATGGC 1520

RESULT 11

US-08-096-043-9
; Sequence 9, Application US/08096043
; Patent No. 5530189
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L
; APPLICANT: Brinkhaus, Friedhelm L
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H
; APPLICANT: Yarger, James G
; APPLICANT: Yen, Huel-Chu B
; TITLE OF INVENTION: Lycopene Biosynthesis in
; TITLE OF INVENTION: Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago

STATE: IL
COUNTRY: USA
ZIP: 60680-0703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
FILING DATE: 22-JUL-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/785,568
FILING DATE: 30-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Galloway, No. 5530189val B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 3128567180
TELEFAX: 3128564972
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1522 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-096-043-9

Query Match 10.1%; Score 708.4; DB 1; Length 1522;
Best Local Similarity 67.6%; Pred. No. 5.6e-184;
Matches 1029; Conservative 0; Mismatches 481; Indels 12; Gaps 2;

QY 3685 ATGAACGCACTTATGATTGGCGCAGGCTTTGGCGGCTGCGGCTGGCGATTCGCGCTG 3744
Db 11 ATGAAAAAACCCGTGTGATTGGCGCAGGCTTTGGTGGCTGCGCTGGCGATTCGCGCTG 70
QY 3745 CAAGCGCGGCGCATACCAACCACTTACTGAGCAGCGCGCAAAACCGCGGCGAGCGCC 3804
Db 71 CAGCGCGCAGGATCCCAACCGTACTGCTGAGCAGCGGGAACAAGCCGCGCGTCCGGGCC 130
QY 3805 TATGTTTGGAGCAGTGGCTTTACCTTGCATCCCGGACCCACGGTGATCACCAGATCCC 3864
Db 131 TACGCTGGCAGTACCGAGGCTTTTACCTTTGACCCCGGCGCAGCGTGATCACCAGATCCT 190
QY 3865 AGCGGCATCGAAGAGTTGTTTCACTGCTGGCAAGAAATGCTCAGCGATTCAGTCAAGCTG 3924
Db 191 ACCGCGCTTGGAGCGCTGTTTCACTTGGCGGCGGAGGATGAGGATTCAGTCAAGCTG 250
QY 3925 ATGCGGTAAACCGCTTCTATCGCTGCTGGGAAGATGCAACACAGCTTGTATACGAC 3984
Db 251 CTGCGGTAAACCGCTTCTATCGCTGCTGGGAGTCCGGGAAGACCCCTCGACTATGCT 310
QY 3985 AATATACGCGCTGCTGGAGCAGATCCCACTGTTCAATCCGCAAGATGTAGAAGGC 4044
Db 311 AACGACAGCTTGGAGCGGAGATTAACCCAGTTCAACCCCGCGAGCTCGAGGCG 370
QY 4045 TATCGTCAATTTCTTGCCTATTCAGCTGAAGATTTAGAGAGGTTATCTGAAGTCCGC 4104
Db 371 TACCGCGCTTCTGCTTACTCCAGCGGCTTATCCAGGAGGATATTTGGCGCTCGC 430
QY 4105 ACGTGGCGGTTCTGCGAGTGGTGACATGCTGGCGCTCGCGCGAGTGGAGCTG 4164
Db 431 AGCGTGGCGGTTCTCTCTTTTCGACATGCTGGCGCGGCGGCGGCTGCTTAAGTCTC 490
QY 4165 CAAGCATGGCGCAGCGCTCTACAGCATGTTGGCGAAATTTATTCAGAGCATCATCTGGT 4224
Db 491 CAGCGGTGGCAGAGGCTTACCGAGTGGTTTTCGCGCTTTATTTAGAGATGAGCATCTGGC 550
QY 4225 CAGCGGTTTTCTTCCATCATCTGTTGGCGGTAATCTTTTGGCAACGTCATCGATC 4284
Db 551 CAGCGCTTCTGTTCTCCCTGCTGTGAGGCGGCAACCCCTTCCACACCTCGTCCATC 610

QY 4285 TATACCTTAATTCATCGCTGAGCGTGAATCGGGGCTGTGGTTTCGCGCGCGGACC 4344
Db 611 TACACCTGATACAGCCCTTGAGCGGAGTGGGGGTCTGGTTCCCTGAGCGCGGACC 670
QY 4345 GCGCGCTGTGTCAGGCGATCGCGGACTGTTCGAGGACTTGGGGGCGGAGCTGTACTG 4404
Db 671 GGGCGCTGTGAACCGCATGTGAAGCTGTTTACCGATCTGGCGGGGAGATCGAATC 730
QY 4405 AATGCGAAGTGAAGCGAGTGAACACGAGCGCATCGAATAGCGGGCTTCAGTTAGAG 4464
Db 731 AACGCCCGGTGGAAGAGCTGTGTGGCGGTAACCGGTAAAGCCAGGTCCGGCTCGCG 790
QY 4465 GCGGACGAGCGCTTCGATCGCGCGCTGTGGCTTCAATGCCGAGTGTGCTACCTAC 4524
Db 791 GATGTGCGATCTTTGACACCGAGCGGTAGGCTCGAAGCTGAGCTGTGTGAACCTAT 850
QY 4525 GACAACTGCTTCGCGACCATCGCTGGCAATGAACGTCGACATCGCTGAAGCGTAAG 4584
Db 851 AAAAGCTGCTCGGACCATACCGGTGGGCGAGAGCGGCGGACGCTGGAGCGCAAG 910
QY 4585 CGCATGAGCACTCGCTGTGTGCTTCTATTTTGGCTGTAATCGAGCGGATGAACAGCTC 4644
Db 911 AGCATGAGCACTCGCTGTGTGCTTCTACTTCGGCTGAACAGCTCAITTCAGCTG 970
QY 4645 GCGCACACACCGCTGTGTGGCGCGGTATCGTGTGATCGATGAGATTTTCAAC 4704
Db 971 GCGCACCATACCATCTGTGTGGTCCCGCTACCGGAGCTGATCGAGAGATCTTTACC 1030
QY 4705 AGCAGCAGCTGGCAGAGATTTTCACTTTACCTGACGCGCCCTGCAGAGCGATCCG 4764
Db 1031 GCGAGCGCTGGCGGATGACTTCTGCTTACCTGCACTGCGCTGGTGACCATCC 1090
QY 4765 TCGTGGCAACCGCGCTGCGGAGCTTTTATGTGTAGCGCGGCTCGCATCTCGGC 4824
Db 1091 TCGCTCGCGCTCCCGCTGCGGAGCTTCTACGTGCTGGCGCGCTGCGCATCTTGC 1150
QY 4825 ACCGCTGACATCGCTGCAACAGAGAGCGCGCTTGGCGATCGAATTTTCTTAT 4884
Db 1151 AACGCGCGCTGGATGCGGCGAGAGGGGCGGAAAGCTGCGGAGCGCATCTTTGACTAC 1210
QY 4885 CTGAGCAGCACTACATCGCGGATTAAGTCAAGCAATAGTGACACACAGAAATGTTTACG 4944
Db 1211 CTGAGAGCGCTATATGCGCGCTGCTGAGCAGCTGTGACCCAGCGATCTTTACC 1270
QY 4945 CGTTTGTATTTGCGGACGCTGCATGCGCATCCATCAGCTCGCGCTTTCGCTGAGCGG 5004
Db 1271 CGGACAGCTTCAACGACGCTTGGATCGGATCTTGGGATCGCTTTTCATCGAGCGCT 1330
QY 5005 ATTTTGCAGCAAGCGCTGTGTCGCGCGATACCGGATCGCGATCAGCAATCTC 5064
Db 1331 TCGTTGACCCAGGCTTGTTCGCGCAACCGCACAGCATTC-----AAACCTC 1381
QY 5065 TATCTGGTGGTCCGCTACGATCCAGCGCGGCGTCCCGGCTGATCGGTTTCGCGC 5124
Db 1382 TACCTGTGGCGGAGTACTACCTGCGCGGCGATCTTGGGCTAGTGGGCTCGCC 1441
QY 5125 AAGGCCACCGCAGCTGATGCTGAGGATCGCGCGAATGAATCGACAGCTTTACTTG 5184
Db 1442 GAAAGCACCGCAGCTGATGATTGAGGATC---TGCAATGAGCAACCGCGCTGCTTG 1498
QY 5185 AGCAAGTAACGCAACCATGGC 5206
Db 1499 ACCACGCCAGCTCGACCATGGC 1520

RESULT 12

US-08-096-623A-9

; Sequence 9, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedrich L.
; APPLICANT: Mukharji, Indrani

; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Hwei-Che B.
; TITLE OF INVENTION: Biosynthesis of Zeaxanthin and
; TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Welsh & Katz, Ltd.
; STREET: 120 S. Riverside Plaza, 22nd Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/096,623A
; FILING DATE: 22-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,061
; FILING DATE: 09-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/662,921
; FILING DATE: 28-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/562,674
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/525,551
; FILING DATE: 18-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/487,613
; FILING DATE: 02-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Gambon, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: AMO-006.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 655-1500
; TELEFAX: (312) 655-1501
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1522 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; POSITION IN GENOME:
; MAP POSITION: -10 to 1512
; UNITS: bp
; US-08-096-623A-9

Query Match 10.1%; Score 708.4; DB 1; Length 1522;
Best Local Similarity 67.6%; Pred. No. 5.6e-184;
Matches 1029; Conservative 0; Mismatches 481; Indels 12; Gaps 2;

QY 3685 ATGAACGCATCTATGCTATGCTTGGCGCAGCTTGGCGGCTGGCGCTGGCGATTCGCTG 3744
Db 11 ATGGAATAAAACCGTTGTGATTTGGCGCAGGCTTTGGTGGCTTGGCGCTGGCGATTCGCTG 70
QY 3745 CAAGCGGCGGCATACCAACACCTTACTCGAGCAGCGCGACAAACCGGCGGAGCGCGC 3804
Db 71 CAGCGGCGAGGATCCCAACCGTACTGCTGAGCAGCGGACACAGCCCGCGGTGGGCC 130
QY 3805 TATGTGTTTGGAGCAGTGGCTTTTACCTTCGATGCGGACCCACCGGTGATCAACCGATCC 3864
Db 131 TACGTCTGGCATGACAGGCTTTTACCTTTGACCGCGGCGCGCGGTGATCAACCGATCCT 190
QY 3865 AGCGCATCGAAGAGTTGTTTACGCTGGCAGGAAATCGCTCAGCGATTCGCTGAGCTG 3924

Db	191	ACCGCGCTTGAGCGCTGTTTACCTTGGCGGACGGCGATCGAGGATTCGTCAAGCTG	250
Qy	3925	ATCCGGTAAACGCCCTTCTATATCGCCTGTCTGGGAAGATGGCAAAACAGCTGTGATTACGAC	3984
Db	251	CTCGCGTAAACCCCTTCTACCGACTCTGCTGGAGTCGCGGAAGACCCCTCGACTATGCT	310
Qy	3985	AATTAATCAGCCGCTGTGGAGCAGCAGATGCCACGTTCAATCCGCAAGATGTAGAAGCC	4044
Db	311	AACGACAGCTTCGAGCTTTGAGGGCGAGATTACCGAGTTCAACCCCGCGAGCTCGAGGCG	370
Qy	4045	TATCGTCAATTTCTTGCTATTACGTGAAGTATTTAGAGAGGGTTATCTGAAAATCTGGC	4104
Db	371	TACCGGCGCTTCTGGCTACTCCAGGCGGTATTCAGAGAGGATATTTGGCGCTCGGC	430
Qy	4105	ACGGTGCCTTTCTGCAGGTGCTGCATGCTCGCGTCCGCCGAGTTGGAGCTGTG	4164
Db	431	AGGTGCGCTTCTCTCTTTTCGCGATGCTCGCGCGCGGCGCGAGCTGTTAAGTTC	490
Qy	4165	CAAGCATGGCGCAGCGCTTACAGCATGGTGGCGAAATTTATTACGACGATCATCTGGT	4224
Db	491	CAGCGTGGCAGAGCGTCTACCACTCGGTTTCGCGTTTATTGAGGATGAGCATCTGGG	550
Qy	4225	CAGCGTTTCTTCCACTCATTTGCTGCTGGCGGTAACTCTTTTCAACGTCATCGATC	4284
Db	551	CAGCGCTTCTGTTTCCACTCCCTGCTGGTATAGCGGCAACCCCTTACCACCTCTGTCATC	610
Qy	4285	TATACCTTAATTCATGCGCTGGAGCGTGAATGGGGGCTGTGTTCTCCGGCGCGGCACC	4344
Db	611	TACACCTGATCAACGCCCTTGAGCGGAGTGGGGGTCTGGTTCTTGAGGCGGCGACC	670
Qy	4345	GGCGCTGTGTCAGGGCATGGCGGACTGTTTCAGGACTTGGCGCGGAGCTGTACTG	4404
Db	671	GGGCGCTGTGTGAACGGCATGGTGAAGCTGTTTACCGATCTGGCGCGGGAGATCGAATC	730
Qy	4405	AATGCCGAAGTGAGCCAGCTGGAACACAGCGGCAATCGCAATTAGCGCGTTCAAGTTAGAG	4464
Db	731	AACGCCGGGTGGAAGAGCTGGTGGCCGATACCGCGTAAGCCAGGTCGGCTCGCG	790
Qy	4465	GGCGGACGACTTCGATGTCGCGCGCTGCGCTTCAATCCGACGCTGGTGCACTTAC	4524
Db	791	GATGTCGATCTTTGACACGACGCGCTAGCCTCGAAACGTCAGCTGGTGTGAACACTAT	850
Qy	4525	GACAACTGCTTCGCCACCATCCGCTGCAATGAACGTGCGCATCGCTGAAGCGTGAAG	4584
Db	851	AAAAAGCTGCTCGGACCATACCGTGGGCGCAGAGCGGGCGCACGCTGTGAGCGCAAG	910
Qy	4585	CGCATGAGCAACTCGCTGTTTGTACTATTTTGGCCTGAAATCAGCGCATGAACAGCTC	4644
Db	911	AGCATGAGCAACTCGCTGTTTGTCTACTTTCGGCTGAAACCGACCTCATTCGCCAGCTG	970
Qy	4645	GGGACACACGCTGTTTGGCGCGGTTATCGTGAAGTTGATCGATGAGATTTTCAAC	4704
Db	971	GGGACCATACCATCTGTTTGGTCCCGCTACCGGAGCTGATCGAGAGATCTTTACC	1030
Qy	4705	AGCAGCCAGCTGGCAGACGATTTTCACTTTACTTCGACGCGCCTCGACGAGCATCCG	4764
Db	1031	GGCAGCGCGCTGGCGGATGACTTCTTCGCTCTACCTGCACCTCGCCCTCGTGAACCGATCCC	1090
Qy	4765	TGCTTGGCACCGCCGGCTGCGGACGCTTTATGTTTAGCGCGGTGCGCATCTCGGC	4824
Db	1091	TCGCTCGCGCCTCCCGGTGGCGCAGCTTCTACGTGCTGGCCCCGGTGGCGCATCTTGGC	1150
Qy	4825	ACCGTGAATCACTGGCAACAGGAAGGACCGCGCTTTCGCGCATGCAATTTTTCCTTAT	4884
Db	1151	AACGCCCGCTGACTGGCGCAGGAGGGCGGAAGCTGCGCACCGCATCTTTGACTAC	1210
Qy	4885	CTGGAGCAGACTACATGCCGGGATTCAGTCAGCAATTAGTGAACACACAGATCTTTACG	4944
Db	1211	CTTGAAGACGCTATATGCCCGGCTGCGTAGCAGCTGGTAGCCACCGGATCTTTACC	1270
Qy	4945	CGTTTGATTTTTCGACACGCTGCATGCCATCACCGCTCGGGCTTTTCGCTGAGCGCG	5004

Db	1271	CGGCAGACTTTCAGACACACGCTGGATCGCGATCTTGGGATCGCTTTTTCATCGAGCGCGCT	1333
Qy	5005	ATTTTCAGCGCAAGCGCTGGTTCCGCCCGGCATAAACCGCGATCCGATATCAGCAATCTC	5064
Db	1331	TCGTTGACCCAGGCTTGTTCCGCCGCAACGCGACACGACATTCTC-----AAACCTC	1381
Qy	5065	TATCTGGTGGTGCCGGTACGCATCCAGCGCGGGCGTGCCCCGGCGTGATCGGTTCCGGCC	5124
Db	1382	TACTGTGGTCCGCGAGTACTCACTCCCTGGCGCGGCATTCTCTGGCGGTAGTGGGCGCTCGCC	1441
Qy	5125	AAGGCCACCGCCAGGCTGATGCTGGAGGATCGCGCGCAATGAATCGACAGCCTTTTACTTG	5184
Db	1442	GAAGCACCGCCAGCCTGATGATTGAGGATC---TGCATGAGCCACACCGCGCTGCTTG	1498
Qy	5185	AGCAAGTAACGCAAAACCATGCC	5206
Db	1499	ACCAGCCACGTCGACCATGCC	1520
RESULT 13			
US-08-660-645A-6			
; Sequence 6, Application US/08660645A			
; Patent No. 6087152			
; GENERAL INFORMATION:			
; APPLICANT: Hohmann, Hans-Peter			
; APPLICANT: Pasamontes, Luis			
; APPLICANT: Tessier, Michel			
; APPLICANT: van Loon, Adolphus			
; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION			
; NUMBER OF SEQUENCES: 47			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Hoffmann-La Roche Inc.			
; STREET: 340 Kingsland Street			
; CITY: Nutley			
; STATE: NJ			
; COUNTRY: USA			
; ZIP: 07110			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/660,645A			
; FILING DATE: 07-JUN-1996			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: EP 9510888.9			
; FILING DATE: 09-JUN-1995			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Pokras, Bruce A.			
; REGISTRATION NUMBER: 32,748			
; REFERENCE/DOCKET NUMBER: RAN 6002/170			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (201) 235-5801			
; TELEFAX: (201) 235-2363			
; INFORMATION FOR SEQ ID NO: 6:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1482 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
; MOLECULE TYPE: DNA (genomic)			
US-08-660-645A-6			

	Query Match	8.8%	Score 619.4;	DB 3;	Length 1482;
	Best Local Similarity	63.7%	Pred. No. 1.5e-159;		
	Matches 941;	Conservative	0;	Mismatches 536;	Indels 0;
QY	3685	ATGAAACGCACATTATGTGATTCGGCGCCAGGCTTTGGCGCGCTTGGCGCGCTGGCGGATTCGCGTG	3744		
nb	1	ATGAGTTCGGCAGTCGTATCGGGCAGGTTTCGGCGGGCTTGGCGCTTGCATCCGCGTG	60		

QY 4765 TCGCTGGACCGCCCGGCTGCGGAGCTTTATGTGTAGCGCGGTCGCCGATCTCGGC 4824
 Db 1081 AGCTTGGCCCGGAGGGGATGTCACGCAATAGTCTTGGCCCGTTCGCGATCTGGGC 1140
 QY 4825 ACCGTGACATCGACTGGCAAGGACCGGCTTGGCGGATCGAATTTTCTTAT 4884
 Db 1141 CGCGCGATGTGATTTGGGAAGCGGCGCTATGCCGAGCGCATCTTCGAGGAA 1200
 QY 4885 CTGGAGCAGCATACATGCGGGATTTACGTACGCAATAGTGACACACAGAAATGTTTACG 4944
 Db 1201 CTGGAGCGCGCGCATATCCGACCTGCGCAAGCACTGACCGTCAGCGCATCTTCAGC 1260
 QY 4945 CCGTTTGAATTTTGGCACACGCTGCAATACGCGCTCGGCTTTCGCTGGAGCGG 5004
 Db 1261 CCGCGCATTTTACGACCAAGCACTGTCGCGCCATACGCGGAGCGCTTTCGCTGGAGCGG 1320
 QY 5005 ATTTTACGCAAGCGCTGTTTCCGCGCGATACCGCGATGCGGATATCAGCAATCTC 5064
 Db 1321 ATCTGACGCAATCCGCTGTTTCCGCGCGATATCCGCGATACCGCGGATCCGCAATCTC 1380
 QY 5065 TATCTGTGTGGTTCGCGGTACGCAATCCGCGCGGCGGCTGCGGCGGTGATCGGTTGCGGC 5124
 Db 1381 TACATCGTGGGCGGCGACGCAATCGGCTGCGGCAATCCGCGGTGCTGCTGGCAGCGCC 1440
 QY 5125 AAGCCACCGCGAGCTGATGCTGGAGGATCGCGCG 5161
 Db 1441 AAGCCACCGCGAGCTGATGCTGCGGACCTGGCGG 1477

RESULT 15
 US-09-546-969-6
 ; Sequence 6, Application US/09546969
 ; Patent No. 6207409
 ; GENERAL INFORMATION:
 ; APPLICANT: Hohmann, Hans-Peter
 ; APPLICANT: Pasamontes, Luis
 ; APPLICANT: Tessier, Michel
 ; APPLICANT: van Loon, Adolphus
 ; TITLE OF INVENTION: FERMENTATIVE CAROTENOID PRODUCTION
 ; NUMBER OF SEQUENCES: 47
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann-La Roche Inc.
 ; STREET: 340 Kingsland Street
 ; CITY: Nutley
 ; STATE: NJ
 ; COUNTRY: USA
 ; ZIP: 07110
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/546,969
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/660,645
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Pokras, Bruce A.
 ; REGISTRATION NUMBER: 32,748
 ; REFERENCE/DOCKET NUMBER: RAN 6002/170
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (201) 235-5801
 ; TELEFAX: (201) 235-2363
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1482 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)

QY 3685 ATGAACACCACTTATGTGATGGGCGAGCTTTTGGCGGCTTGGCGCTGCGATTCGCGTG 3744
 Db 1 ATGAGTTTCGGCCATCGTCAATCGGCGAGTTTCGGCGGCTTGGCGCTTGGCGATTCGCGTG 60
 QY 3745 CAAGCGGCGGATACCAACACCTTACTTCAGACGAGCGGCAACACCGGCGGAGCGGCC 3804
 Db 61 CAATCGGCGGATCGCGACCAACCATCTGTCAGGCGCGCGCAACCGCGGCGGCGCGGCC 120
 QY 3805 TATGTGTTTGGAGACAGTGGCTTTTACCTTCGATGCGGAGCGGCGGCGGATCACCGATCCC 3864
 Db 121 TATGTCTGGAACGATCAGGCGCACTCTTCGATGAGCGCGCGGCGGATGACCGACCCC 180
 QY 3865 AGCGCCATCGAAGAGTTGTTCAGCTGGCAGGAAATCTGCTCAGGATGATGCTGAGCTG 3924
 Db 181 GACAGCTCGGAGAGCTGTGGGCGCTCAGCGGCGCAACCGATGAGCGGTGACGCTG 240
 QY 3925 ATGCGGTTAAACCGCTTCTATCGCTGCTGTGGGAGATGGCAACAGCTTGTATACGAC 3984
 Db 241 CTGCGGCTCTCGCCCTTCTACCGGCTGACATGGCGGAGCGCGCGAGCTTCGATACGTG 300
 QY 3985 AATAATCAGCGCTGCTGAGCAGCAGATCGCCACGTTCAATTCGCAAGATGTAGAAGGC 4044
 Db 301 AAGCAGCAGCAGCTGATCGCCAGCTCGCTCTTCAATCCCGCGGATGTGATGCG 360
 QY 4045 TATCGTCAATTTCTGCTTATTCAGTGAAGTATTTAGAGGCTTATCTGAACTCGGC 4104
 Db 361 TATCGCGCTTCCACGATTTACGCGGAGAGGTCTATCGCGAGGGGTATCTGAAGCTGGG 420
 QY 4105 ACGGTCCGCTTCTGAGGCTGCTGACATGCTGCGGCTGCGCGCGGAGTTGGAGCTCTG 4164
 Db 421 ACCACGCTTCTGAGCTGGCGAGATGCTGAACGCGCGCGCGCTGATGCGCTG 480
 QY 4165 CAAGCATGCGCGAGCTCTACAGCATGCTGGGGAATTTATTCAGCAGCATCTCTGGT 4224
 Db 481 CAGGCAATACGCTCGCTGACAGCATGCTGGGCGCTTTCATCCAGGACCGCATCTGGGG 540
 QY 4225 CAGCGCTTCTTCCACTCATCTGCTGGGCGGTAATCTTTTCAACGTCATCGATC 4284
 Db 541 CAGCGCTTCTGCTTCCACAGCTGCTGGTGGCGGGAACCGTTCGACCAAGCTGATC 600
 QY 4285 TATACCTTAATTCATGCTGGAGCGGTGAATGGGCGGTGTGTTTCGCGCGGCGGCGACC 4344
 Db 601 TATGCGCTGATCCATGCTGGAACGCGCGCGGCTGCTGTTTCGCGCAAGCGCGCAC 660
 QY 4345 GCGCGCTGCTGAGGCGCATGCGGCTGCTGAGGACTTGGCGCGGAGCTGTTACTG 4404
 Db 661 AACCAGCTGCTGCGGCGCATGCTGCGCTTTCGAGCGCTTTCGCGCGACGCTGCTGCTG 720
 QY 4405 AATCCGAGTGAGCCAGCTGGAACACGCGGCAATCGCAATAGCGGCGTTCAGTTAGAG 4464
 Db 721 AATCCGCGCTCAGCGGATCGACACCGAGGCGGATCGCGCCAGCGGCTGCGCTGCTG 780
 QY 4465 GCGGAGCAGCTTTCGATGCGCGCTGCTGCGCTTCAATCGCGGCTGCTGATACCTAC 4524
 Db 781 GACGCGGCGAGTTTTCGCGCGGATACGCTGCGGCAACCGGCGGATGATGACAGCTAT 840
 QY 4525 GACAACTGCTTCCGCGCATTCGCTGCGCAATGAACTGCGCATCGCTGGAAGCGTGAAG 4584
 Db 841 CGGACCTTCTGGGCGCATACCGCGCGGCGGCGACCAAGCGCGGATCTTGAACCGGCG 900
 QY 4585 CGCATGAGCAACTCGCTGTTTGTACTTATTTTGGCGGCTGATGATGAGATTTTCAAC 4644
 Db 901 CGTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 QY 4645 GCGGACACACGCTGTTTGGCGCGGCTTATCGTGAATGATGATGATGATGATGATGATGAT 4704
 Db 961 GCGGACACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY 4705 AGCAGCAGCTGCGGAGAGATTTTCTACTTTTCACTTTCAGCGCGCTTTCAGCGGATCCG 4764
 Db 1021 GGGCCACGCTTGGCGGAGATTTTCTCGATGATGATGATGATGATGATGATGATGATGATG 1080

US-09-546-969-6

Query Match 8.8%; Score 619.4; DB 3; Length 1482;
Best Local Similarity 63.7%; Pred. No. 1.5e-159;
Matches 941; Conservative 0; Mismatches 536; Indels 0; Gaps 0;

```
QY 3685 ATGMAACGACATATGTGATTCGCGCAGAGCTTTGGCGGCTTGGCGCTGCGCTGCGATTCGCCCTG 3744
Db 1 ATGAGTTCCGCATCGTCTGATCGCGCAGAGTTTCGGCGGGCTTGCCTTCCCATTCGCCCTG 60

QY 3745 CAAGCGCGGGGATACCAACACCTTACTCGAGCAGCGCAAAACCGGGCGGAGCGGCC 3804
Db 61 CAATCGCGCGCATCGCAGCACCATCTGTCGAGGCCCGCGACAGCCCGCGCGCGGCC 120

QY 3805 TATGTGTTTGGAGACAGTGGCTTTACTCTTCGATGCGCGACCCACACGATGATCAACGATCCC 3864
Db 121 TATGTCTGGAACGATCAGGCGCACAGCTCTTCGATGACAGGCCCGACAGCTCGTGACCGACCCC 180

QY 3865 AGCGCCATCGAAGATTTGTCAGGCTGGCAGGAATATCTCAGCGATTACGTCGAGCTG 3924
Db 181 GACAGCTCGGAGAGCTGTGGGCCCTCAGCGGCCAACCGATGGAGCGTGACGTGACGCTG 240

QY 3925 ATGCCGCTAACGCCCTTCTATCGCCTGTGTGCGGAAGATGCGAAACAGCTTGAATACGAC 3984
Db 241 CTGCGGCTCTCGCCCTTCTACCGGCTGACATGGCGGACGCGCGCAGCTTCGATACGCTG 300

QY 3985 AATAATCAGCGCTGCTGAGCAGCAGATGCGCAAGTTCAATCCGCAAGATGTAGAAGCC 4044
Db 301 AACGACGACGAGCTGATCGGCCAGGTTCGCTCTCAATCCCGCCGATTCGATGCG 360

QY 4045 TATCGTCAATTTCTTCCCTTATTCAGTGAAGTATTAGAGAGGTTATCTGAACTCGGC 4104
Db 361 TATCGCGCTTCCACGATTACGCGAGAGGTTATTCGCGAGGGGTATCTGAAGCTGGGG 420

QY 4105 AGCGTCCCTTCTGCGAGTGGTGATGCTGCGGTCGCGCGCGCAGTTGGAGCCTG 4164
Db 421 ACCACGCCCTTCTTGAAGTGGGCGCAGATGCTGAACGCGCGCGCGGCTGTAGCGCTG 480

QY 4165 CAAGCATGCGCAGCGCTTACAGCATGTTGGGGAATTTATTCAGGACGATCATCTGGCT 4224
Db 481 CAGGATACCGCTCGGTCACAGCATGTTGGGCGCTTCATCCAGGACCGCATCTGCGG 540

QY 4225 CAGGCTTTTCTTCCATCATGCTGTTGGGGTAACTCTTTGCAACGTCATCGATC 4284
Db 541 CAGGCTTCTGCTTCACACGCTGCTGCTGGCGGGAACCCGTTTTCGACAGCTCGATC 600

QY 4285 TATACCTTAATTCATGCGCTGAGCGTGAATGGGCGTGTGTTTCCGCGCGCGCAC 4344
Db 601 TATGCGCTGATCCATGCGCTGGNACGCGCGCGCGGCTGTGTTCCGCAAGGCGCGCAC 560

QY 4345 GCGCGCTGCTGGCAGGCGATGCGCGCATGTTTCGAGGACTTGGGCGGCGAGTGTACTG 4404
Db 661 AACCACTGCTGGCGGCGATGCTGCGCTTTCGAGCGTCTTGGCGGCGACGCTGCTGCTG 720

QY 4405 AATGCCAAGTGAAGCAGCTGGAACACGCGGCMATCGCATTAGCGCGCTTCAGTTAGAG 4464
Db 721 AATGCCCGGCTACGCGGATGACACCCGAGGCGGATCGCGCCACGCGGCGTCACTGCTG 780

QY 4465 GCGGACGACGCTTCGATGCGCGCTGTGGCTTCCAAATGCGGACGCTGGTGATACCTAC 4524
Db 781 GACGGGCGGAGTTGTGGCGCGGATACGCTGGCCGACGACGCGGATGACAGCTAT 840

QY 4525 GACAACTGCTTGGCCACCATCGCTGGCAATGAACGTTGCGACATCGCTGAAGCGTAAG 4584
Db 841 CGGACCTGCTGGGCGCATACCGCGCGCGGCGCACCAAGGCGCGGATCTGAAACCGGCG 900

QY 4585 CGCATGAGCACTCGCTGTTTGTACTTATTTTGGCTGAATCAGCGCGCATGAACAGCTC 4644
Db 901 CGCTGCTGATGCTGCTGTTTGTGCTGCTATTTGGCTGTCTCAAGCGCGCGGAGACCTG 960

QY 4645 GCGCACACACGCTGTTTGGCGCGCTTATCGTGAGTTGATCGATCAGATTTTTCAC 4704
Db 961 GCCACCAACAGCTATCTTGGCGCGCGCTACAGGGGCTGTGNAACGAGATCTTCAAC 1020
```

Search completed: January 25, 2005, 18:28:48

Job time : 509 secs

```
QY 4705 AGCAGCCAGCTGGCAGACGATTTTTCATCTTACCTGACGCGCCCTTCGACGACGATCCG 4764
Db 1021 GGGCCACGCTCGCCGACGATTTTCTCGATGTATCTGCAATTCGCTCGCTGACCCGATCCC 1080

QY 4765 TCGCTGGCACCGCCCGGCTGCGGACGCTTTTATGTGTTAGCGCGCTGCGCATCTCGGC 4824
Db 1081 AGCCTGGCCCCCGAGGGGATGTCACGCAATACGTCCTTGGCGCCGCTTCGCACTCGGGC 1140

QY 4825 ACCGCTGACATCACTGCGCAACAGAAAGGACCGCGCTTTCGCGGATCGAAATTTTGTCTAT 4884
Db 1141 CGCGCGGATGTCGATTGGGAAGCCGAGGCCCGGCTATGCGAGCGCATCTTCGAGGAA 1200

QY 4885 CTGGAGCAGCATACATGCGCGGATTAAGTCAAGCAATTAAGTGACACACAGAAATGTTTACG 4944
Db 1201 CTGGAGCGCGCGCCATCCCGACCTGCGCAAGCACCTGACCGTCAGCCGCGCATCTTCAGC 1260

QY 4945 CGGTTTGTATTTTCGCGACACGCTGCAATGCCCATCACGGCTCGCGCTTTTTCGCTGGAGCCG 5004
Db 1261 CCGCGGATTTTCAGCACCGAACTGTCGGCCCATCACGGCAGCGCTTCTCGGTCGAGCCG 1320

QY 5005 ATTTTGACGCAAGCGCCTGTTTCGCGCGCATAAACCGGATGCGGATATCAGCAATCTC 5064
Db 1321 ATCCTGACGCAATCCGCTGTTTCGCGCGCATAAACCGGATGCGGATCCCGAACTTC 1380

QY 5065 TATCTGTTGGTGGTCCGCTACGATCCAGCGCGCGGCTGCCCGCGCTGATTCGCTTCGCGC 5124
Db 1381 TACATCGTGGGGCGGCGACGATCCGGTTCGCGGCTCCCGGCTGCTGTTGGCAGCGCC 1440

QY 5125 AAGGCCACGCGCAGGCTGATGCTGGAGGATCGCGCG 5161
Db 1441 AAGGCCACGCGCAGGCTGATGCTGGACCTGGGCGG 1477
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

QM nucleic - nucleic search, using sw model

Run on: January 25, 2005, 10:29:07 ; Search time 3350 Seconds
(without alignments)
12004.619 Million cell updates/sec

Title: US-10-808-979-18
Perfect score: 6999
Sequence: 1 gcgatggcaatggaaatgt.....gacgttttagacatgaata 6999

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptoddata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptoddata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptoddata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptoddata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptoddata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptoddata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptoddata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptoddata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptoddata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptoddata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptoddata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptoddata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptoddata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptoddata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptoddata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptoddata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptoddata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptoddata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptoddata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptoddata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptoddata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6999	100.0	6999	18	US-10-808-979-18
2	4326.2	61.8	8675	18	US-10-808-979-18
3	2619.6	37.4	8609	18	US-10-735-019-28
4	2619.6	37.4	8609	18	US-10-734-778-40
5	2619.6	37.4	8609	18	US-10-734-936-37
6	2619.6	37.4	8609	18	US-10-735-442-64
7	2138.6	30.6	8814	18	US-10-808-807-18
8	2135.2	30.5	12753	16	US-10-041-018-19
9	1618.8	23.1	7494	17	US-10-474-536-45
10	1618.8	23.1	8547	17	US-10-474-536-46
11	1482	21.2	1482	18	US-10-808-979-7
12	1278	18.3	1278	18	US-10-808-979-3

13	1167	16.7	1167	18	US-10-808-979-5	Sequence 5, Appli
14	1085.2	15.5	1482	18	US-10-804-677-7	Sequence 7, Appli
c 15	1035.4	14.8	8625	10	US-09-920-923-1	Sequence 1, Appli
c 16	1035.4	14.8	8625	16	US-10-695-980-1	Sequence 1, Appli
c 17	1035.4	14.8	11233	10	US-09-920-923-27	Sequence 27, Appli
c 18	1035.4	14.8	11233	16	US-10-695-980-27	Sequence 27, Appli
19	930	13.3	930	18	US-10-808-979-9	Sequence 9, Appli
20	906	12.9	906	18	US-10-808-979-1	Sequence 1, Appli
21	875.4	12.5	1278	18	US-10-804-677-3	Sequence 3, Appli
22	861.4	12.3	1479	18	US-10-808-807-7	Sequence 7, Appli
23	808.6	11.6	1479	10	US-09-941-947A-31	Sequence 31, Appli
24	808.6	11.6	1479	15	US-10-218-118-7	Sequence 7, Appli
25	808.6	11.6	1479	16	US-10-363-567-31	Sequence 31, Appli
26	808.6	11.6	1479	16	US-10-466-656-7	Sequence 7, Appli
27	808.6	11.6	1479	18	US-10-735-019-7	Sequence 7, Appli
28	808.6	11.6	1479	18	US-10-734-778-7	Sequence 7, Appli
29	808.6	11.6	1479	18	US-10-735-442-7	Sequence 7, Appli
30	808.6	11.6	1479	18	US-10-848-307-7	Sequence 7, Appli
31	695	9.9	1167	18	US-10-804-677-5	Sequence 5, Appli
32	692.6	9.9	1521	14	US-10-166-037-5	Sequence 5, Appli
33	630.8	9.0	930	18	US-10-804-677-9	Sequence 9, Appli
34	629.2	9.0	906	18	US-10-804-677-1	Sequence 1, Appli
35	619.4	8.8	1482	9	US-09-547-267-6	Sequence 6, Appli
36	535.6	7.7	1473	15	US-10-369-493-23921	Sequence 23921, A
c 37	534	7.6	534	18	US-10-808-979-11	Sequence 11, Appli
38	510.6	7.3	1518	15	US-10-369-493-41517	Sequence 41517, A
39	480.8	6.9	1494	15	US-10-369-493-34266	Sequence 34266, A
40	449.2	6.4	1296	10	US-09-941-947A-27	Sequence 27, Appli
41	449.2	6.4	1296	15	US-10-218-118-3	Sequence 3, Appli
42	449.2	6.4	1296	16	US-10-363-567-27	Sequence 27, Appli
43	449.2	6.4	1296	16	US-10-466-656-1	Sequence 1, Appli
44	449.2	6.4	1296	18	US-10-735-019-3	Sequence 3, Appli
45	449.2	6.4	1296	18	US-10-734-778-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-808-979-18
; Sequence 18, Application US/10808979
; Publication No. US20040268439A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2360 US NA
; CURRENT APPLICATION NUMBER: US/10/808,979
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US 60/471,904
; PRIOR FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 6999
; TYPE: DNA
; ORGANISM: Enterobacteriaceae strain DC260
US-10-808-979-18

Query Match	100.0%	Score 6999;	DB 18;	Length 6999;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6999;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	GGCATGGCAATGGAAATGTTCCGTGAGCTTTTCGTTCTTAACGGCGCATCAGCGGTGAA	60	
Db	1	GGCATGGCAATGGAAATGTTCCGTGAGCTTTTCGTTCTTAACGGCGCATCAGCGGTGAA	60	
QY	61	TAATCAGCACGCGAGGTTGGTATTGCGTCATTTTATGTCATTAGCGGTATTTTATTTAC	120	
Db	61	TAATCAGCACGCGAGGTTGGTATTGCGTCATTTTATGTCATTAGCGGTATTTTATTTAC	120	

QY	121	CACACGCTTAACGAGTATTTATGTTAGGGGTTGCTGTGTTACTTATTTCCCTAATAAGA	180	1201	TTTAAATCAACAGCTAGCGATATTTCACTGAGCGCGCTCAGCGGTGGCCCACTTTGCGG	1260
Db	121	CACACGCTTAACGAGTATTTATGTTAGGGGTTGCTGTGTTACTTATTTCCCTAATAAGA	180	1201	TTTAAATCAACAGCTAGCGATATTTCACTGAGCGCGCTCAGCGGTGGCCCACTTTGCGG	1260
QY	181	TCACCATAGCATTTATTAACATATTTACCTGGTGGCATGAATACGACACCTTCAAAAGT	240	1261	TGATCGCGCGCGCTCTACAGCACTTTTCAACGCGTTGAGGGGTTAGCACAACCGCTGC	1320
Db	181	TCACCATAGCATTTATTAACATATTTACCTGGTGGCATGAATACGACACCTTCAAAAGT	240	1261	TGATCGCGCGCGCTCTACAGCACTTTTCAACGCGTTGAGGGGTTAGCACAACCGCTGC	1320
QY	241	CAAGTCCCTCGCTGGCGAACTCACCTTAGCGAGTCTACGGTTAAATCAAAAAGCATAAAA	300	1321	TGCGCGCGCGCATCCGATCACATTCATCCAGCAAGCCGATGCCCGACTTTGCTTTAGCG	1380
Db	241	CAAGTCCCTCGCTGGCGAACTCACCTTAGCGAGTCTACGGTTAAATCAAAAAGCATAAAA	300	1321	TGCGCGCGCGCATCCGATCACATTCATCCAGCAAGCCGATGCCCGACTTTGCTTTAGCG	1380
QY	301	TTTACCAACCAATGATAGCATATTAACACCAACCATGTCACACACAGCATCAGACA	360	1381	ACGAAACGATTCGATTTTGTTCGGTGGGCAACAGACGATCTCTCGGTTGCTGGCGC	1440
Db	301	TTTACCAACCAATGATAGCATATTAACACCAACCATGTCACACACAGCATCAGACA	360	1381	ACGAAACGATTCGATTTTGTTCGGTGGGCAACAGACGATCTCTCGGTTGCTGGCGC	1440
QY	361	ACGAACTCTCTCAGCTGCAGCAATTTTACAGGCGCATCTTGAACATTTACTGCTGCC	420	1441	CGGTGTCATCGGCTGGCTCGCGGGCGGCTGTCGCTGTTTTCGCTGATCGACGATC	1500
Db	361	ACGAACTCTCTCAGCTGCAGCAATTTTACAGGCGCATCTTGAACATTTACTGCTGCC	420	1441	CGGTGTCATCGGCTGGCTCGCGGGCGGCTGTCGCTGTTTTCGCTGATCGACGATC	1500
QY	421	GGACAGCAAGCATCGCTGCGTGGCGGATGCGTGGCGGAACGCTGGCGGAGGGCAAA	480	1501	TCGCTCTCTCACCGATATGCTGTGCCGCAACTGCTGCGGTACTGAAAGCATTTGAACA	1560
Db	421	GGACAGCAAGCATCGCTGCGTGGCGGATGCGTGGCGGAACGCTGGCGGAGGGCAAA	480	1501	TCGCTCTCTCACCGATATGCTGTGCCGCAACTGCTGCGGTACTGAAAGCATTTGAACA	1560
QY	481	CGTATTCGTCCTTTATTAATCTGCTGGCAGCGCGCATGATGGTTGCGAGCTGACGCA	540	1561	TCGATGCGGTGATCGCCGACGAAATGGAAGCGGGCGGATTTGGTTCGCTGAAAGCGCTGC	1620
Db	481	CGTATTCGTCCTTTATTAATCTGCTGGCAGCGCGCATGATGGTTGCGAGCTGACGCA	540	1561	TCGATGCGGTGATCGCCGACGAAATGGAAGCGGGCGGATTTGGTTCGCTGAAAGCGCTGC	1620
QY	541	AATGGCGTTCTCGATCTCGCTGTGAGTGGAAATGGTGACCGCGCATTCGTGATTCG	600	1621	ATCTGCCGTTTGTTCGCTGGCTGCGCTTTCGCGGTCAATCGTGAAGCGGGATTCGCG	1680
Db	541	AATGGCGTTCTCGATCTCGCTGTGAGTGGAAATGGTGACCGCGCATTCGTGATTCG	600	1621	ATCTGCCGTTTGTTCGCTGGCTGCGCTTTCGCGGTCAATCGTGAAGCGGGATTCGCG	1680
QY	601	GATGACATTCCTCGATGATACGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	660	1681	TTGCGGTGATGCCCTTCGCTTTTGCAACGATGAAAGCGCTGAAACGTTTTCAGGCCA	1740
Db	601	GATGACATTCCTCGATGATACGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	660	1681	TTGCGGTGATGCCCTTCGCTTTTGCAACGATGAAAGCGCTGAAACGTTTTCAGGCCA	1740
QY	661	GAATTTGGTGAACAGTGGGATTTCTCGCGCATCGCTGCTTAGCGCGCATTTGAA	720	1741	GCAGGATATCTATGATCGCATCATGCTGTCACGCGAGTGTATCTCTCAAAACACGCGC	1800
Db	661	GAATTTGGTGAACAGTGGGATTTCTCGCGCATCGCTGCTTAGCGCGCATTTGAA	720	1741	GCAGGATATCTATGATCGCATCATGCTGTCACGCGAGTGTATCTCTCAAAACACGCGC	1800
QY	721	GTGATGCTTACGACCGGTTTTCGCTGCGCATATTAATCTGAAGCGATTCGTAATC	780	1801	GGCGGTTTAAATTTGACGAGCGCGCGATTTACATCAGTGTCTGCGCGCTGGCACAAA	1860
Db	721	GTGATGCTTACGACCGGTTTTCGCTGCGCATATTAATCTGAAGCGATTCGTAATC	780	1801	GGCGGTTTAAATTTGACGAGCGCGCGATTTACATCAGTGTCTGCGCGCTGGCACAAA	1860
QY	781	TCGCTGCTGCTGCGCTGCGAGGCTTAGTCAAGGCAATTCAGGATCTGACACGCGC	840	1861	TCAGCAGATGGTGGCGGCTTTGATTTTCCAGCTCAGCAACTGCGCGCTGCTATCACG	1920
Db	781	TCGCTGCTGCTGCGCTGCGAGGCTTAGTCAAGGCAATTCAGGATCTGACACGCGC	840	1861	TCAGCAGATGGTGGCGGCTTTGATTTTCCAGCTCAGCAACTGCGCGCTGCTATCACG	1920
QY	841	ACGACAGCGCGAGCGCGAGCGATCGCATGACCAAGAACTGAAACACCGCGTGTG	900	1921	CCGTGGGCGCACTCGCGCGCGCGGTTTCTCTGCGCGCTCCATGCGCGCTGGCCAGCGC	1980
Db	841	ACGACAGCGCGAGCGCGAGCGATCGCATGACCAAGAACTGAAACACCGCGTGTG	900	1921	CCGTGGGCGCACTCGCGCGCGCGGTTTCTCTGCGCGCTCCATGCGCGCTGGCCAGCGC	1980
QY	901	TTTCGCGCGCATGCAATTTGGCGGATTTGCGCTGAGCTTTCAGCGAGTGGCGCAA	960	1981	TGCGTCAAGCGCGGTTTATGCTGCTGGGTAAGTCAAGGCGCATCGCTTCGGGCTGT	2040
Db	901	TTTCGCGCGCATGCAATTTGGCGGATTTGCGCTGAGCTTTCAGCGAGTGGCGCAA	960	1981	TGCGTCAAGCGCGGTTTATGCTGCTGGGTAAGTCAAGGCGCATCGCTTCGGGCTGT	2040
QY	961	AGACTTACCTTTCGCGCGAGGATTTGGGCGAGCGTTTCACTGCTGCAAGCACTCGCC	1020	2041	TTCTGCTATCGCGCAGCGCTGCGCGAGTGGCTATCGCTGATGATCGCCCATTTGTG	2100
Db	961	AGACTTACCTTTCGCGCGAGGATTTGGGCGAGCGTTTCACTGCTGCAAGCACTCGCC	1020	2041	TTCTGCTATCGCGCAGCGCTGCGCGAGTGGCTATCGCTGATGATCGCCCATTTGTG	2100
QY	1021	GACGTTTGAACACACCGGTAAGATGTGACAGGATCAGGCAATCCAGCTGGTA	1080	2101	GGGGATTAACCGCGCAACAGCGCATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGG	2160
Db	1021	GACGTTTGAACACACCGGTAAGATGTGACAGGATCAGGCAATCCAGCTGGTA	1080	2101	GGGGATTAACCGCGCAACAGCGCATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGG	2160
QY	1081	CAGATGCTCGTGTGAGCGCGGCAACCTGCTGCGCATCACCTGCGCAGCGAGAT	1140	2161	ATTTGCTGATCAGCGCGCGCTTACAGCGCGCTGTTTATCTACTATGCGCGGT	2220
Db	1081	CAGATGCTCGTGTGAGCGCGGCAACCTGCTGCGCATCACCTGCGCAGCGAGAT	1140	2161	ATTTGCTGATCAGCGCGCGCTTACAGCGCGCTGTTTATCTACTATGCGCGGT	2220
QY	1141	GCACACCTTTCGCTGCGCTGCGCATCGCGGATCGCCCACTGCGGCTGCGGATTCG	1200	2221	TAAACAGCGCGCTGGAAGCATCTGGAATCGGTACCGCATGCTGCGCTGCGGATTCG	2280
Db	1141	GCACACCTTTCGCTGCGCTGCGCATCGCGGATCGCCCACTGCGGCTGCGGATTCG	1200	2221	TAAACAGCGCGCTGGAAGCATCTGGAATCGGTACCGCATGCTGCGCTGCGGATTCG	2280
				2281	TTGATCAGCGCGCGCTGGCGCGCATTTGATGCGATGACGTTGCTGCGCGCGCATCAC	2340

Db 2281 |||||TTGATCAGCCGCGTGGCGCGCAATGAGTGGCATGACGTGTGCTCGCGCATCAC 2340
Qy 2341 GCTTTAGCCGCTGTTCACTAACTGGAGCAGCATCTGCAACAGAGTGTCTGACCGCAGATCGTT 2400
Db 2341 GCTTTAGCCGCTGTTCACTAACTGGAGCAGCATCTGCAACAGAGTGTCTGACCGCAGATCGTT 2400
Qy 2401 AGCCGCTACGAGTGTAGCGATTCAGCGCAGTGTGAGCGCGAGCGCGTGTGCGAGCGTG 2460
Db 2401 AGCCGCTACGAGTGTAGCGATTCAGCGCAGTGTGAGCGCGAGCGCGTGTGCGAGCGTG 2460
Qy 2461 CCGCGAGACATCGTGGAGAGCGCTGTGCGAGCAAGTGTGCTGTGCGAGAGCGACCT 2520
Db 2461 CCGCGAGACATCGTGGAGAGCGCTGTGCGAGCAAGTGTGCTGTGCGAGAGCGACCT 2520
Qy 2521 GATGCGCAGCAATACGATGTGATTTTGGTTCGCTGTGACCTGGCGAATGGCTTGATTCG 2580
Db 2521 GATGCGCAGCAATACGATGTGATTTTGGTTCGCTGTGACCTGGCGAATGGCTTGATTCG 2580
Qy 2581 GCTGGCTGTGCTCAATTCGAGCACAACCTGAAATGCCTGTGCGAATGGCTTGATTCG 2640
Db 2581 GCTGGCTGTGCTCAATTCGAGCACAACCTGAAATGCCTGTGCGAATGGCTTGATTCG 2640
Qy 2641 TCCGCGAGCAATCATACCTGTGCTGTTTCACTCAGCGCATCTCAGCGCGAACAATTCG 2700
Db 2641 TCCGCGAGCAATCATACCTGTGCTGTTTCACTCAGCGCATCTCAGCGCGAACAATTCG 2700
Qy 2701 CTGGCTGCACACCGCTGATTTACCGTGGTGGTTCAGGTATCAGGTGCGTTCCTGCGCT 2760
Db 2701 CTGGCTGCACACCGCTGATTTACCGTGGTGGTTCAGGTATCAGGTGCGTTCCTGCGCT 2760
Qy 2761 GCGCGCAATCTGGAGCGGGAATATGTTTCATTCGATCAGCGGATTTTTCGCGCATCT 2820
Db 2761 GCGCGCAATCTGGAGCGGGAATATGTTTCATTCGATCAGCGGATTTTTCGCGCATCT 2820
Qy 2821 TTACGCGGAGTGGTGCAGTCTGGACAAACACAGCGGTACACAGGTAAACCCAC 2880
Db 2821 TTACGCGGAGTGGTGCAGTCTGGACAAACACAGCGGTACACAGGTAAACCCAC 2880
Qy 2881 GCAGGTGACCTGGCGGATGGCGGTGAACTTGTGCGCAAGTGGTGAATGATGGTGGCGG 2940
Db 2881 GCAGGTGACCTGGCGGATGGCGGTGAACTTGTGCGCAAGTGGTGAATGATGGTGGCGG 2940
Qy 2941 CTGAGCGCGACCGCACATCTGAGCTGGGTTATCAGGTGTTTCTGGACAAGAGTGGCA 3000
Db 2941 CTGAGCGCGACCGCACATCTGAGCTGGGTTATCAGGTGTTTCTGGACAAGAGTGGCA 3000
Qy 3001 GCTGGCGAGCGCAGCGCTGAGCAGCGGATCTGTGATGGATGCCACCGTCGATCAGCA 3060
Db 3001 GCTGGCGAGCGCAGCGCTGAGCAGCGGATCTGTGATGGATGCCACCGTCGATCAGCA 3060
Qy 3061 AGCGGGTTATCGTTTGTCTACACGCTGCGCTCAGCGCGGATCGGCTATTGATTGAAGA 3120
Db 3061 AGCGGGTTATCGTTTGTCTACACGCTGCGCTCAGCGCGGATCGGCTATTGATTGAAGA 3120
Qy 3121 TACCCATTACGTTAACCGCCGCTGCGGAGAACACCGCTGCTGAGCAATCGCGCA 3180
Db 3121 TACCCATTACGTTAACCGCCGCTGCGGAGAACACCGCTGCTGAGCAATCGCGCA 3180
Qy 3181 CTATGCCAATCAGCAAGGTGAGCGTGTGATGAGTGTGCTGCGTGAAGAGCACGGCATATT 3240
Db 3181 CTATGCCAATCAGCAAGGTGAGCGTGTGATGAGTGTGCTGCGTGAAGAGCACGGCATATT 3240
Qy 3241 ACCGATTACCTGAGCGGCAACATCGATCGATTCTGGCAACAGCAGCGCGGCAACGCTG 3300
Db 3241 ACCGATTACCTGAGCGGCAACATCGATCGATTCTGGCAACAGCAGCGCGGCAACGCTG 3300
Qy 3301 CAGCGCCTGCGCGCGGCTGTTTCAATGCCACACCGGTTACTCTTCCGCTGCGCGCT 3360
Db 3301 CAGCGCCTGCGCGCGGCTGTTTCAATGCCACACCGGTTACTCTTCCGCTGCGCGCT 3360
Qy 3361 GCGCTAGCGGAGTTGGTAGCAGCGCTGTGCGCACCGATGCGCTCAGCTCAGCGCAACA 3420

Db 3361 GCGCTAGCGGAGTTGGTAGCAGCGCTGTGCGCCACCGATGCCCTCAGCCTCAGCAACA 3420
Qy 3421 TATCGAACGCTTTGCGCGCTCAGCAGTGGCGGAAACAGAGATTTTCCGCTGCTAAACCG 3480
Db 3421 TATCGAACGCTTTGCGCGCTCAGCAGTGGCGGAAACAGAGATTTTCCGCTGCTAAACCG 3480
Qy 3481 CATGCTGTTTGGCGCGTAAAGCCGAGCAGCGCTGGCGGTGATGCAACGTTTTCACCG 3540
Db 3481 CATGCTGTTTGGCGCGTAAAGCCGAGCAGCGCTGGCGGTGATGCAACGTTTTCACCG 3540
Qy 3541 GCTCATGCGCGGTTAAATAGCCGCTTTTACGCCGGCAACTGCGCTCGCGGATAAAC 3600
Db 3541 GCTCATGCGCGGTTAAATAGCCGCTTTTACGCCGGCAACTGCGCTCGCGGATAAAC 3600
Qy 3601 GCGGATTTCTGTGCGCAAGCCGCTGCCATCGGTGAAGCGCTGCGCGCTGTGTA 3660
Db 3601 GCGGATTTCTGTGCGCAAGCCGCTGCCATCGGTGAAGCGCTGCGCGCTGTGTA 3660
Qy 3661 TTCTGTGCAACCCAGGGAAGAAAATGAAACGCACTTATGATTTGGCGCAGGCTTTGCG 3720
Db 3661 TTCTGTGCAACCCAGGGAAGAAAATGAAACGCACTTATGATTTGGCGCAGGCTTTGCG 3720
Qy 3721 GGCCTGGCGCTGGCGATTCGCTGCAAGCGGCGCATACCAACCACTTACTCGAGCAG 3780
Db 3721 GGCCTGGCGCTGGCGATTCGCTGCAAGCGGCGCATACCAACCACTTACTCGAGCAG 3780
Qy 3781 GCGCAAAACCGGCGGCGCGCTTATGTTTGGAGCAGTGGCTTTTACCTTCGATGCC 3840
Db 3781 GCGCAAAACCGGCGGCGCGCTTATGTTTGGAGCAGTGGCTTTTACCTTCGATGCC 3840
Qy 3841 GGACCCACGCTGATCAGCATCCAGCGCATCGAAGAGTTGTTTCACTGGCGAGGAAA 3900
Db 3841 GGACCCACGCTGATCAGCATCCAGCGCATCGAAGAGTTGTTTCACTGGCGAGGAAA 3900
Qy 3901 TCGCTCAGCGATTACGTGAGCTGATGCGGTAAACCGCTTCTATCGCTGCTGGGAA 3960
Db 3901 TCGCTCAGCGATTACGTGAGCTGATGCGGTAAACCGCTTCTATCGCTGCTGGGAA 3960
Qy 3961 GATGCAAAACAGCTTTGATTACGACAAATATCAGCGCGCTGCTGGAGCAGATCGCACG 4020
Db 3961 GATGCAAAACAGCTTTGATTACGACAAATATCAGCGCGCTGCTGGAGCAGATCGCACG 4020
Qy 4021 TTCAATTCGCAAGATGTAGAGGCTATCGTCAATTTCTTGCCCTATTACGTGAATTTT 4080
Db 4021 TTCAATTCGCAAGATGTAGAGGCTATCGTCAATTTCTTGCCCTATTACGTGAATTTT 4080
Qy 4081 AGAGAGGTTATCTCAAACTCGGCAAGTGGCGTTCGAGTGGTGCATGCTGCGC 4140
Db 4081 AGAGAGGTTATCTCAAACTCGGCAAGTGGCGTTCGAGTGGTGCATGCTGCGC 4140
Qy 4141 GTGCGCGCGAGTTGGGAGCTCTGCAAGCATGGCGCAGCGTCTACAGCATGGTGGCGAAA 4200
Db 4141 GTGCGCGCGAGTTGGGAGCTCTGCAAGCATGGCGCAGCGTCTACAGCATGGTGGCGAAA 4200
Qy 4201 TTTATTCAGGACGATCATCTGCGTCAGGGTTTTTCTTCCATCTCATTTGCTGGCGGCT 4260
Db 4201 TTTATTCAGGACGATCATCTGCGTCAGGGTTTTTCTTCCATCTCATTTGCTGGCGGCT 4260
Qy 4261 AATCCCTTTGCAACGTCATCTATACCTTAAATTTCTGCGCTGGAGCGTGAATGGGCG 4320
Db 4261 AATCCCTTTGCAACGTCATCTATACCTTAAATTTCTGCGCTGGAGCGTGAATGGGCG 4320
Qy 4321 GTGCTGTTTTCGCGCGGCGCACCGCGCGCTGGTGCAGGGCATGGCGGACTGTTTCGAG 4380
Db 4321 GTGCTGTTTTCGCGCGGCGCACCGCGCGCTGGTGCAGGGCATGGCGGACTGTTTCGAG 4380
Qy 4381 GACTTTGGCGCGAGCTGTTACTGAATGCGAAGTGAGCCAGCTGGAAACAGCGGCAAT 4440
Db 4381 GACTTTGGCGCGAGCTGTTACTGAATGCGAAGTGAGCCAGCTGGAAACAGCGGCAAT 4440
Qy 4441 GCGATTAGCGGCTTACGTTAGAGGCGGAGCAGCGCTTCGATGCGCGCTGCTGGCCTCC 4500
Db 4441 GCGATTAGCGGCTTACGTTAGAGGCGGAGCAGCGCTTCGATGCGCGCTGCTGGCCTCC 4500

QY 4501 AATGCCGAGCTGGTGCATACCTACGACAAACCTGCTTCGCCACCATCCGCTGGCAATGAAA 4560
DB |||||
QY 4501 AATGCCGAGCTGGTGCATACCTACGACAAACCTGCTTCGCCACCATCCGCTGGCAATGAAA 4560
DB |||||
QY 4561 CQTGCGACATCGCTCAAGCGTAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGC 4620
DB |||||
QY 4561 CQTGCGACATCGCTGAGGGTAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGC 4620
DB |||||
QY 4621 CTGAATCAGCCGATGAACAGCTCCGCGACCAACACCGTCTGTTTGGCCCGGGTTATCGT 4680
DB |||||
QY 4621 CTGAATCAGCCGATGAACAGCTCCGCGACCAACACCGTCTGTTTGGCCCGGGTTATCGT 4680
DB |||||
QY 4681 GAGTTGATCGATGAGATTTTCAACAGCAGCGAGCTGGCAGAGATTTTTCACTTTACCTG 4740
DB |||||
QY 4681 GAGTTGATCGATGAGATTTTCAACAGCAGCGAGCTGGCAGAGATTTTTCACTTTACCTG 4740
DB |||||
QY 4741 CACGCGCCCTGCAGCAGCGATCCGTCGCTGGCACCGCCCGCTGGCGAGCTTTTATGTG 4800
DB |||||
QY 4741 CACGCGCCCTGCAGCAGCGATCCGTCGCTGGCACCGCCCGCTGGCGAGCTTTTATGTG 4800
DB |||||
QY 4801 TTAGCGCCGGTCCGATCTCGGACCGCTGACATCGACTGGCAACAGGAGGACCGGC 4860
DB |||||
QY 4801 TTAGCGCCGGTCCGATCTCGGACCGCTGACATCGACTGGCAACAGGAGGACCGGC 4860
DB |||||
QY 4861 TTGCGGATCGAATTTTTCGTTTATCTGGAGCAGCACTACATGCCGGGATTAAGTCAAGCAA 4920
DB |||||
QY 4861 TTGCGGATCGAATTTTTCGTTTATCTGGAGCAGCACTACATGCCGGGATTAAGTCAAGCAA 4920
DB |||||
QY 4921 TTAGTGACACACAGAAATTTTACGCCGTTTGAATTTTTCGCGACACGCTGATGCCATCAC 4980
DB |||||
QY 4921 TTAGTGACACACAGAAATTTTACGCCGTTTGAATTTTTCGCGACACGCTGATGCCATCAC 4980
DB |||||
QY 4981 GGCCTCGGCTTTTCGCTGAGCGGATTTTGAAGCAAGCGCTGTTTTCGCCCGCATTAAC 5040
DB |||||
QY 4981 GGCCTCGGCTTTTCGCTGAGCGGATTTTGAAGCAAGCGCTGTTTTCGCCCGCATTAAC 5040
DB |||||
QY 5041 CGCGATGCCGATATCAGCAATCTCTATCTGGTGGGTGCGGTCAGCGATCCAGCGCGGGC 5100
DB |||||
QY 5041 CGCGATGCCGATATCAGCAATCTCTATCTGGTGGGTGCGGTCAGCGATCCAGCGCGGGC 5100
DB |||||
QY 5101 GTGCCCGGCTGATCGGTTTCGGCAAGGCCACCGCAGCGCTGATGCTGAGGATCGCGCC 5160
DB |||||
QY 5101 GTGCCCGGCTGATCGGTTTCGGCAAGGCCACCGCAGCGCTGATGCTGAGGATCGCGCC 5160
DB |||||
QY 5161 GAATGAATCGACAGCCCTTTACTTTGACCAAGTAACGCAACCACTTGGCGTGGCTCGAAGA 5220
DB |||||
QY 5161 GAATGAATCGACAGCCCTTTACTTTGACCAAGTAACGCAACCACTTGGCGTGGCTCGAAGA 5220
DB |||||
QY 5221 GTTTTCGCCACCGCCGCAAGCTGTTTGTGACACCGACCGCCCGCAGCACGCTGATGCTGT 5280
DB |||||
QY 5221 GTTTTCGCCACCGCCGCAAGCTGTTTGTGACACCGACCGCCCGCAGCACGCTGATGCTGT 5280
DB |||||
QY 5281 ATCGCTGGTGTCTGCTACCTGGGATGATGATGATGGGCAAGCGCTGGCGAGGCGGCA 5340
DB |||||
QY 5281 ATCGCTGGTGTCTGCTACCTGGGATGATGATGATGGGCAAGCGCTGGCGAGGCGGCA 5340
DB |||||
QY 5341 CGCAGCATCGCTCGAAGACCGCAGCAGCTATGAGCATCTGCAAAATTTGAAACCCGCC 5400
DB |||||
QY 5341 CGCAGCATCGCTCGAAGACCGCAGCAGCTATGAGCATCTGCAAAATTTGAAACCCGCC 5400
DB |||||
QY 5401 GCGCCTACAGCGCGCGCACATGATGAACCGCGGTTTAGGGGTTTCAAGGAAGTGGCA 5460
DB |||||
QY 5461 TCATTACAGCTGCCGCAACACTGGCTTTTCATCTGGAAGGCTTCGCTATGATG 5520
DB |||||
QY 5461 TCATTACAGCTGCCGCAACACTGGCTTTTCATCTGGAAGGCTTCGCTATGATG 5520
DB |||||
QY 5521 CAGCAACGAACTTACGCGAGCTTCGATGACACGCTGCGTTACTGCTATCACGTCGCGG 5580
DB |||||
QY 5521 CAGCAACGAACTTACGCGAGCTTCGATGACACGCTGCGTTACTGCTATCACGTCGCGG 5580
DB |||||

QY 5581 GCCTGTGCTGGTTTGAATGCGCGCGGTAAATGGCGCTGGCGAGCAAGACGGTCTCTGATC 5640
DB |||||
QY 5581 GCCTGTGCTGGTTTGAATGCGCGCGGTAAATGGCGCTGGCGAGCAAGACGGTCTCTGATC 5640
DB |||||
QY 5641 ACGCTGCGGATTTAGGATCGCGCTTCAGCTCACTAACTTGGCGCGGACATTTGTAGAAG 5700
DB |||||
QY 5641 ACGCTGCGGATTTAGGATCGCGCTTCAGCTCACTAACTTGGCGCGGACATTTGTAGAAG 5700
DB |||||
QY 5701 ATGCCGAAATGTTGCTGCTATCTGCGCAATCTTGGCTCGATCAGGCGGGATTTACGCG 5760
DB |||||
QY 5701 ATGCCGAAATGTTGCTGCTATCTGCGCAATCTTGGCTCGATCAGGCGGGATTTACGCG 5760
DB |||||
QY 5761 CCGATACGCTGACTGACCCGCAACATCGTGAGCGCTCGCTCACTGCGAGCGGTTTAG 5820
DB |||||
QY 5761 CCGATACGCTGACTGACCCGCAACATCGTGAGCGCTCGCTCACTGCGAGCGGTTTAG 5820
DB |||||
QY 5821 TGGCGGAGGCGGAAACCTTATCTACCTGGCGGATCCGGTTTACCGGTTTACCGCTGC 5880
DB |||||
QY 5821 TGGCGGAGGCGGAAACCTTATCTACCTGGCGGATCCGGTTTACCGGTTTACCGCTGC 5880
DB |||||
QY 5881 GCTCGGCTGGGCGCATCGCTACCGCTCGCGCGCTTTTATCGGAAATTTGGCGTCAAGATT 5940
DB |||||
QY 5881 GCTCGGCTGGGCGCATCGCTACCGCTCGCGCGCTTTTATCGGAAATTTGGCGTCAAGATT 5940
DB |||||
QY 5941 AGCAACGCGGTGTCACGCTGGATTCACGCGAGCGCACAGTAAAGGTGAAACCTGG 6000
DB |||||
QY 5941 AGCAACGCGGTGTCACGCTGGATTCACGCGAGCGCACAGTAAAGGTGAAACCTGG 6000
DB |||||
QY 6001 CGCTGTGTTGAAAGGGGAGGTTTGGCGATCACTTCGCGTGTCTCTGCTCTGAAACCGC 6060
DB |||||
QY 6001 CGCTGTGTTGAAAGGGGAGGTTTGGCGATCACTTCGCGTGTCTCTGCTCTGAAACCGC 6060
DB |||||
QY 6061 GTCCGGCTGGTCTCTGTCAGCGCTCTCTGTTGATTTTACGTTACGTTCCGTCAGCGTCAGCGT 6120
DB |||||
QY 6061 GTCCGGCTGGTCTCTGTCAGCGCTCTCTGTTGATTTTACGTTCCGTCAGCGTCAGCGT 6120
DB |||||
QY 6121 GGCCTGACGCTTATTCAGCGGTGGCGGTAGAGAAACCAACACAGCAGCGCTTCACG 6180
DB |||||
QY 6121 GGCCTGACGCTTATTCAGCGGTGGCGGTAGAGAAACCAACACAGCAGCGCTTCACG 6180
DB |||||
QY 6181 CCGCGCACCCATGATGATCGCGTCGCGCATGTATAAGCGCTTAAAGATAGCCTTTGCG 6240
DB |||||
QY 6181 CCGCGCACCCATGATGATCGCGTCGCGCATGTATAAGCGCTTAAAGATAGCCTTTGCG 6240
DB |||||
QY 6241 CCGGATATAGCGGAAACCGGCTGATGATGACACAGGCGCTGCGACCATGAGTAGAG 6300
DB |||||
QY 6241 CCGGATATAGCGGAAACCGGCTGATGATGACACAGGCGCTGCGACCATGAGTAGAG 6300
DB |||||
QY 6301 CCGCGCTGACGTCGTCATTTCCGCGACCAATCCACTGACGCGGCGCACATGCTTGCACACC 6360
DB |||||
QY 6301 CCGCGCTGACGTCGTCATTTCCGCGACCAATCCACTGACGCGGCGCACATGCTTGCACACC 6360
DB |||||
QY 6361 GACATAAATCAGCAAAATCGCCAGTACCGCAAAACCAACCGCATGATGTTGAGCTC 6420
DB |||||
QY 6361 GACATAAATCAGCAAAATCGCCAGTACCGCAAAACCAACCGCATGATGTTGAGCTC 6420
DB |||||
QY 6421 AAACTTACCGCTGCTGCTGTCATGTTGCGGACAGATGCCAGCCCATCCCAACCGTGCAT 6480
DB |||||
QY 6421 AAACTTACCGCTGCTGCTGTCATGTTGCGGACAGATGCCAGCCCATCCCAACCGTGCAT 6480
DB |||||
QY 6481 GATGTATTTATGCGACAGCGCGCTTACGATTTTCATCAACCGGTTGCGCAACAGAT 6540
DB |||||
QY 6481 GATGTATTTATGCGACAGCGCGCTTACGATTTTCATCAACCGGTTGCGCAACAGAT 6540
DB |||||
QY 6541 AAGCAGCTTCCATAACAGAGATTTGTCGCTCATTTTGGGAAAGGAGTACTAAAGG 6600
DB |||||
QY 6541 AAGCAGCTTCCATAACAGAGATTTGTCGCTCATTTTGGGAAAGGAGTACTAAAGG 6600
DB |||||
QY 6601 TGGACGCGGATGAGTGAGGCGCAAGTTTACCATGTTTAGAAATTTTAAAGTCCATAA 6660
DB |||||
QY 6601 TGGACGCGGATGAGTGAGGCGCAAGTTTACCATGTTTAGAAATTTTAAAGTCCATAA 6660
DB |||||
QY 6661 CACGTTATGACGCTGCTATTCAGAAAGCGCAGATTTTACATATCTCAACACCTTATC 6720
DB |||||

```
Db 6661 CACGTTATGAACGCTGCATTTGAGAAAGCGCAGATTTTCACATATCTCCACCACTTATC 6720
Qy 6721 AATACACGTTGTTAACTACATGCGGGATTTATGCTTTCTACAGCCGTAAGACAAAAAAA 6780
Db 6721 AATACACGTTGTTAACTACATGCGGGATTTATGCTTTCTACAGCCGTAAGACAAAAAAA 6780
Qy 6781 CTGTGAGTGTGACACTTGAACCTGCTCTACTGAGCAAGCCAGAGAGGAGGCGGCTCAATT 6840
Db 6781 CTGTGAGTGTGACACTTGAACCTGCTCTACTGAGCAAGCCAGAGAGGAGGCGGCTCAATT 6840
Qy 6841 TATCGCCATCTATCCAAAGCTTTGCAACATGAAATTCGACGACTGCGACGACGAGAGAT 6900
Db 6841 TATCGCCATCTATCCAAAGCTTTGCAACATGAAATTCGACGACTGCGACGAGAGAT 6900
Qy 6901 GGAAGCGTGAACACAGTGAAGGTTTGCAGGAACTCAATCGCATACCGAAGAGCAGGTT 6960
Db 6901 GGAAGCGTGAACACAGTGAAGGTTTGCAGGAACTCAATCGCATACCGAAGAGCAGGTT 6960
Qy 6961 TATTGTGCGATGAATACAGGAGCTTTTACATGCAATA 6999
Db 6961 TATTGTGCGATGAATACAGGAGCTTTTACATGCAATA 6999
```

RESULT 2

```
US-10-804-677-18
; Sequence 18, Application US/10804677
; Publication No. US20040224383A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2346 US NA
; CURRENT APPLICATION NUMBER: US/10/804,677
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US 60/468,596
; PRIOR FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; TYPE: DNA
; ORGANISM: Pectobacterium cypripedii DC416
US-10-804-677-18
```

```
Query Match 61.8%; Score 4326.2; DB 18; Length 8675;
Best Local Similarity 80.1%; Pred. No. 0;
Matches 5122; Conservative 0; Mismatches 1268; Indels 5; Gaps 3;

Qy 236 AAGTCAAGTCCCTCGCTGGGAACTCACCTTACGAGTCTAGGTTAATC-AAAAGCA 294
Db 1119 AAGTCTGCTGCTCGCTGGGAACTCCCTTTACCGGGTCTACGGTTAATTGAAAAGCA 1178
Qy 295 TAAATAATTTTCAACCAACATGATAGCATTATGACCAACCATGTCGACACCAACAGACAT 354
Db 1179 CAAGATTTTAACTAACCATGGAAGCGCTATGACCGCCCATGTCGATACCAAGCAAGC 1238
Qy 355 CAGACAAGCGAACTCTCTCAGCTGCGAGCAAAATTTTACAGGCGCATCTTGAAATTTACTG 414
Db 1239 CAGGAAAGCGATCTCTCTCAGTTGCATCAGCATTCGCAATTCAGGCCCCATCTTGAACATTATTG 1298
Qy 415 CTGCGCGGACAGCAAGCATCGCTGCGTGGCGGATGCTGCGGCAACGCTGGGGCAG 474
Db 1299 CTGCGCGGACAGCAAGCATCGCTGCGGCGGCAACGCTGGGGCAG 1358
Qy 475 GCGAAACGTTATCGTCTTTTACTGCTGGCAGCGCGATATGGTTGCGAGCTG 534
Db 1359 GCGAAACGTTATCGTCTTTTACTGCTGGCAGCGCGATATGGTTGCGAGCTG 1418
```

```
Qy 535 AGCGAAATGGCGTTCTCGATCTCCCTGTGTCAGTGGAAATGTCGACGCGGATCGCTG 594
Db 1419 GCGCAGAGGCGATCTCTTGTGCTGTGCGGTGCAAAATGTCGACGCTCCTCAGT 1478
Qy 595 ATTCTGGATGACATTCCTCGATGATAAGCGCAGATGCGTCGTGTCCTCCCTACCGTG 654
Db 1479 ATCTCTGACGACATTCATGATTAACGCCCGGATGCGAGCTGGGGGCCCGGCAATC 1538
Qy 655 CATCGCAATTTTGTGTAATAAGTGGCGATTTCTCGCGCGCATCGCGTGTCTTAAAGCGGCA 714
Db 1539 CACTGTGAATATGGGAAAGCGTGGCGATCTCTGGCAGCGGTGCGCTACTCAGCGCGCC 1598
Qy 715 TTTGAAGTGAATGCGCATTCGACCCCGTTTCCCTGCGCATACATAAATCTGAAGCGATTGT 774
Db 1599 TTTGAGGTGATTTGCGCTCTGCGCGGCTCTGCCAGCAAGCAAAAGCGAAGCCATTGCG 1658
Qy 775 GAACTCTCGCTCGCTCGCGCTGCGAGGCTTTAGTGAAGGCAATTCAGGATCTGCAC 834
Db 1659 GAGCTCTCTCTGCGCTGGGCTGCGAGGACTGGTTCAGGTCAGTTCAGGATCTGCAT 1718
Qy 835 GACGCGACGACAGCGCGGAGCGATGCGCATGACCAACGAATCGAAACCAAGC 894
Db 1719 GACGCGCAGACAGCGCGAGTCCGGAAGCCATCACCTGACCAATGAATGAAACCAAGC 1778
Qy 895 GTGCTGTTTTCGCGCCAGCTGCAAAATGGCGGCGATGCGCTGACGCTTCACGCGAGGTG 954
Db 1779 GTCTGTTTTCGCGCCAGCTGCGAGATGGCGCGATTTGCGGCGATGCGCTGAGTGAAGT 1838
Qy 955 CGGCAAGAGCTTTAGCTTCTTGGCCAGGATTTGGCGCAGGCGTTTCACTGCTCGACGAC 1014
Db 1839 CGTCAGCGTTTAAAGCTATTTTGGCAGGATTTAGTTCAGGCTTTTCAAGTCTTGGACGAC 1898
Qy 1015 CTGCGCGAGCTTGAACACACACCGGTAAAGATGTGCACAGGATCAGGCGCAATCCAGC 1074
Db 1899 CTGCGCGATGGCTCTTAAGCACACACCGCAAGGACTGTCTATCAGATCAGGCGCAATCCAGC 1958
Qy 1075 TGTGTACAGATGCTCGTGTGACGCGCGGGAAGCTGCGCTGCGCGATCAGCTGCGGACG 1134
Db 1959 TGTGTGACAGATGCTGCGGCGCGGAGGCTGAGCGCTGCTGCGGACCATCTAAAGCAGC 2018
Qy 1135 GCAGATGCACACTTTGCTGCGCTGCGCATCGCGCATCGCCACTCGCCAAATATATGAC 1194
Db 2019 GCGGATGCACACTTTGCTGCGCTGCGCATCGCGGTGTCGCCACCGCTCAATATATGAC 2078
Qy 1195 GCGCTGTTTAAATCAACAGCTAGCGATTTCAACTGAGCGCGCTCAGCGGTTGGGCACT 1254
Db 2079 GCGCTGTTTAAATCAACAGCTAGCGATTTCAACTGAGCGCGCTCATACTATGGGCACT 2138
Qy 1255 TTGCGGTGATCGCGCGCGCTCTACAGCACTTTTCAAGCGGTTGAGGCGTTAGCAAAA 1314
Db 2139 TTGCGGTGATTTGCGCGCGCTCTACAGCACTTTTCAAGCGGCTGGGCGCAAA 2198
Qy 1315 GCTGCTGCGCGCGCGCATCGCATCATCTTCAAGCAAGCGCATGCGCGGCTTTCG 1374
Db 2199 GCTGCTGCGCGCGCGCATCGCATCATCTTATCAAGCAAGTATGATGACGACCTTTCG 2258
Qy 1375 TTGCGAGCAAGCACTGATTTTGTGCGCTGCGCAACAGACGATCTTTCGCGGTTGCG 1434
Db 2259 TGAGCGAGCGAGCGCTGCTTTGCGCGCTGCGGAGCGCGCATCTTTCGCGGATCGC 2318
Qy 1435 TGGCGCGCTGTTGATGCGCTGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1494
Db 2319 TCTCAGCGAACTCAGGCGGCTGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2378
Qy 1495 ACGATCTGCGCTGCTGCGCGGATGCTGCTGCGGAACTGCTGCGGCTGCTGCTGCTGCT 1554
Db 2379 ACGATCTGCGCGAGCAACCGGATGCTGCTGCGGAACTGCTGCGGCTGCTGCTGCTGCT 2438
Qy 1555 TGAACATGATGCGCTGATGCGCGGAGCAAAATGGAAGCGCGGCGGATGCTGCTGCTGCTG 1614
Db 2439 TCGAGGTGATGCGCTGATGCTGCGGATCAATGGAAGCGGCTGCTGCTGCTGCTGCTGCT 2498
Qy 1615 CGCTGATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1674
```


Db 6879 AAAAGCTGGCGTCTGCTGTAAGAGTGGCGGCTGCGCTTACTTTCGCGCCTTGTCTCATC 6938
QY 6052 CTGAACCGCGTCTGGGTGCTGTGTGGAGCGTCTCTGTTGATTTTACGTCTGGTACGCTG 6111
Db 6939 CCGAGCGCGTCTGCTGGGCTGTGTGGAGCGTCTGGTGTGACACAGCGCCATGGCGGTG 6998
QY 6112 GCGACGCGGTGCTGAGCTTATTCAGCGGTGGCGGTAGAGAAACCAAAACGACACGCA 6171
Db 6999 GCGACGCGTGGCTGCGCTGAGCTTGTGCAACGCTGGGCGTTAAAGAAAGCAAGAAACGCA 7058
QY 6172 GCGTTTACCGCGCGCACCGCATGATGATGCGGTGGCGCATATGATAAGCGCTTAAGATA 6231
Db 7059 GCGTTTCCCGTCCCGCACCGCATGATGATGCGGTGGCGCATATGATAAGCGCTTACGATA 7118
QY 6232 GCGTTTGGCGGGATATACGGAAACCGCGAGGTGATGATGACACGCGCATGCTGACCAT 6291
Db 7119 GCGTTTGGCGTGAATATAGCAAAACGCGCAGCGTGTGATGGACGAGCGCTCATGACCAT 7178
QY 6292 GAAGTAGCGCGCGCTAGCTCGTCAATTCGCGCACCAATCCACTGACGCGGCCACATGCC 6351
Db 7179 AATAATAGAGAGCGCCATAGTGTCTATGCGAGCGCCATCCACTGAGCGGCCAGAGGCC 7238
QY 6352 TTGCAACCGCATATAATCAGCACAAATGCGCGAGTACCGCAACACACCGCATAAAGATC 6411
Db 7239 GTTGACACCCAGCCAAATCAGGCCAATCGAACACGCGCAACACACCGCATACAGGTC 7298
QY 6412 GTTGAGCTCAAACTTACCGCTGTGGGTTTCATGTTGCGACAGATGCCAGCGCCATCCCA 6471
Db 7299 GTTGAGCTCAAACTTGTCTCTCATGTGTTTCATGATGCGCAAAATGCCAGCGCCATCCCA 7358
QY 6472 ACCGTGCATGATGATTTATGCGACAGCGCGCTACGATTTCCATCACCACCGCGTTGC 6531
Db 7359 GCCATGATATGATTTATGCGACAGCGCGCGGAGTCTCCATCAGTATCAAGTAGC 7418
QY 6532 CAAACAGATTAAGCATGTTCCATTAACAGAGCATGTTGTCCTCATTTGTGGAAGAGGAG 6591
Db 7419 CAGCAGGATAAGCGCATTCATTAACAGAGCATCAATGTTGTTCCATTTTGGCAAGTAGAG 7478
QY 6592 TACTAAAGTGGACGCGGATGAGTGGCGAAG 6626
Db 7479 TA-TAAAGGTGGACGTGGATAGCGAAAGCGCAAG 7512

RESULT 3

US-10-735-019-28
; Sequence 28, Application US/10735019
; Publication No. US20040191863A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Suh, Wonchul
; TITLE OF INVENTION: Mutations Affecting Plasmid Copy Number
; FILE REFERENCE: CL2029 US NA
; CURRENT APPLICATION NUMBER: US/10/735,019
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/434973
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 28
; LENGTH: 8609
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pPCB15
US-10-735-019-28
Query Match 37.4%; Score 2619.6; DB 18; Length 8609;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;
QY 373 CAGCTGCAGCAATTTTACAGGCGCATCTTGAACATTTACTGCTCGCGGACAGCAAGC 432

Db 1627 CAGTTGCTGGCTGATATCGATAGCGCTTGTGATCAGTTACTGCGCGTTTCAAGGTGACGGG 1686
QY 433 GATCCGCTGGCTGGCGGATGCGGTGCGGAACGCTGGCGAGGGCAACGATTTTCGTCT 492
Db 1687 GATTTGTGGGTGCGCGATGCGTGAAGGACGCTGGACCGGGCAACGATTTTCGTCCG 1746
QY 493 TTATTTACTGCTGCTGGCAGCGCGATATGAGTTTGGAGCTGACGCCAAATTTGGGTTCTC 552
Db 1747 ATGCTGCTGTTTAAACAGCGCGGATCTTGGCTGTGGATCACTCAAGGGGATTAAG 1806
QY 553 GATTCGCTGCTGAGTGAATGATGCGCGCATGCTGATCTGATCTGATGATGATGATGATG 612
Db 1807 GATTTAGCTGCGCGTGAATGATGCTGCTGCTGCTGATTTCTGATGATGATGATGATG 1866
QY 613 TCGATGATTAACGCGGATGCGGTGCTGCTGCGCTTACCGCTACCGTGCATCGGAAATTTGGTGA 672
Db 1867 TGCATGAGACGATGCGCGATGCGTGGGGGCGTCCCACCATTCACGCGATGACGTTGAA 1926
QY 673 AACGTGGCGATTTCTGCGCGCATGCGCTGCTTACCGCGCATTTTGAAGTGAATTTGCCATT 732
Db 1927 CATGTGGCGATTTCTGGGGCGGTGCTTTACTCAGCAAAAGCGTTTGGGGTGAATTTGCCGAG 1986
QY 733 GCACCGGTTTGGCTGCGCATACATTAATCTGAAGGATTTGCTGAATCTCGCTGCGCTG 792
Db 1987 GCTGAAGTCTGACGCGCATAGCCAAACTCGCGCGGTGTCGAGCTGCTCCTGCTGCTG 2046
QY 793 GCGCTGACAGGCTTAGTGAAGGCAATTCAGAGATCTGACGAGCGGACGACGAGCGCGC 852
Db 2047 GCGATGAGGCTGTGTTTACGCGCGATTTTAAAGACCTCTCGGAAGGCGATTAACCGCGC 2106
QY 853 AGCGCGGAGGATGCGCATACCAACGAACTGAAACAGCGGTGCTGTTTCGCGCGCAAG 912
Db 2107 AGCGCGGATGCGCATCTTAACCAATCAGTTTAAACAGCAGCGCTGTTTTCGCGCGTCA 2166
QY 913 CTGCAATTTGGCGGATTTGCGCTGACGCTTACCGCAGGTGCGGCAAGACTTTAGCTTC 972
Db 2167 AGCAAAATGGCGTCCATTTGCGGCAACGCGTCTGCGAAGCGCGTGAACCTGCTGCTG 2226
QY 973 TTGCGCCAGGATTTGGGCGAGCGCTTCAACTGCTCGACGACCTCGCGCAGGTTCGAAA 1032
Db 2227 TTCTGCTCGATCTCGGCGAGCGCTTTCAGTTGCTTGAAGCTTTCACGATGCGATGACC 2286
QY 1033 CACACCGTGAAGATGTGACACGATCAGGGGAAATCCAGCTGCTGTTACAGATGCTCGT 1092
Db 2287 GATACCGGCAAGACATCAATCAGGATGAGTAAATCAACGCTGCTCAATTTTATTTAGGC 2346
QY 1093 GCTGACGCGCGGACGCTGCGCTGCGCATCACTGCGCAGCGAGATGACACCTTGGC 1152
Db 2347 TCAGCGCGGTGCAAGAACGCTGCGACAGATTTGCGCTGCGCAGCTGAGACACCTTTCC 2406
QY 1153 TGCGCTGCGCATCGCGCATCGCCACTCGCCAA---TATATGCAAGCGCTGTTTAAATCAA 1209
Db 2407 GCGGATGCGCAAAACGCGCATTTCCACACCACTTTTATTCAGGCTGCTGTTGACAAA 2466
QY 1210 CAGCTAGCGATTTTCACTGAGCGCGCTCAGCGGTGGCCACTTTTGGGTTGATCGCG 1269
Db 2467 AAATCTGCTGCGCTCAGTTTAAAGATGCTGCA-----TGAGCCATTTTTCGCGTATCGGCAC 2521
QY 1270 GCGCGCTCTACAGCCACTTTTCAAGCGTTTTCAGGCGTTTACGACAAACGCTGCTGGCGCGCG 1329
Db 2522 GCGCCCTTTTTCAGCCATGTTTCGCGCTCTGCAAAACCTTGTCTCAGGAATTTAGTGGCGCG 2581
QY 1330 GCCATCGCATCATTTTCAAGCAAGCGGATGCGCGCATCTTGTGTTTACGCGACGAAACGCA 1389
Db 2582 GTCATGCTGTACGTTTTCAGCAACATGACTGCAAAAGCGCTGTTTAAACGCGGACGATGATA 2641
QY 1390 TCGATTTTGTGCGCTGCGCAACAGACGCTCTCGCGCTGCTGCGCGCTGCTGCGCGCGGTTGTC 1449
Db 2642 TCGGATTCAGACCGCTGCGCATGCAAAACGATCTCTCCCGTTTCTTATCGCACTGCTGCTG 2701
QY 1450 ATCGCTGCTGCGCTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509

Db 2702 ACCTGGCGCGCACCACTCGACCCCTCGATGTTAGACTGATCAATGAATGGCAGTA 2761
Qy 1510 GCACCGATATGTCGCCCGCAACTGCTCGCGTACTGAAGCATGTAACATCGATGCGC 1569
Db 2762 CCAGCGATATGCTTTGGCGGAATCGCCCGCGCTTTTCATGCGTTGAGATAGAGGCG 2821
Qy 1570 TGATCGCGACGAATGGAGCGGCGGCGGATGCTCGCTGAAGCGCTGCATCTGCCGT 1629
Db 2822 TGATCGTTGATCAATATGAGCGCGGAGGTGCGATAGTCGCAAGACGTCAGGTCTGCCGT 2881
Qy 1630 TTGTTTCGGTGGCTCGCGCTTTCGGCTCAATCGTGAAGCCGGGATTCGGTTGCGGTGA 1689
Db 2882 TTGTTTCGGTGGCTCGCGCTCGCGCTCAACCGCGAACCAGGTTTGCTCTGCGCGTGA 2941
Qy 1690 TGCCCTTCGGTTTGGCAGAGATGAAGAGCGCTGAAGCGTTTTCAGGCGCAGCAGGATA 1749
Db 2942 TGCCCTTCGAGTACGGCACCGAGCATGCGGCTCGGGAACGCTATACCAACGAGGAAAAA 3001
Qy 1750 TCTATGATCGCATCATGCTGCTACGCGGAGTGCATCTCAAAACAGCGCGGGGTITA 1809
Db 3002 TTATGACTGGCTGATGCGAGTCAAGATCGGTGATCGCGCATCATGATGCGAGATGG 3061
Qy 1810 APTTACCGAGCGCGCGGATTTACATCAGTGCCTGTCGCGCTGGCAGCAAAATACGCCAGA 1869
Db 3062 GTTTAGCCCCGGTGAAAAAATGTCATCATTTGTTTCTCCACTGGCAGCAAAATACGCCAGT 3121
Qy 1870 TGGTCCCGGCTTTGATTTTCACTGACAACTGCCCGCTGCTATCAAGCGGCGGCGC 1929
Db 3122 TGATCCCGAACTGGATTTTCCCGCAAGCGCTGCCAGACTGCTTTTCATGCGGTGGAC 3181
Qy 1930 CACTCCGCG- - -CCCGGTTTCTCTGCGCGCTCCATGCGGCCCTGGCCAGGCGTGCCTC 1986
Db 3182 GCTTACGGCAACCCAGGGGAGCGCGGGGTCACTAACTTCTTATTTTCGTCGCCGACA 3241
Qy 1987 AGCCGGTGGTTATGCTCGCTGGGTAGCTCAAGGCGCATCGCTTCGCGCTGTTTCTGC 2046
Db 3242 AACCCGGTATTTTGGCTCGCTGGGCAACCTGACGGAGATCGTTATGGCTGTTTCAGGA 3301
Qy 2047 ATCTGGCGAGCGTGGCCCGCAGCTGCGGCTATCGCTGGTATGCCCATTTGTGGGGAT 2106
Db 3302 CCATCGCAAGCCCTGCAAGAGTGGATGCGAGTTACTGTTGGCACTGTTGGCGGCC 3361
Qy 2107 TAAACGCCGACAGAGCATGAGCTGGAGCTCGCTGGCGCGCTGGGTGACGATTCG 2166
Db 3362 TCTCAGCAGCAGGCAAGTGAATGAGCTGGCCCGGGCGGAGCAATCAGGTGTGGATTTG 3421
Qy 2167 TCGATCAGCGCGAGCCCTACAGCAGCGCAGCTGTTTATCACTCATGCGCGGTTAAACA 2226
Db 3422 CGGATCAATCCGACGACTTTTCAGAGGACAGTTGACATCAATGATGGTGGGATGAATA 3481
Qy 2227 GCGCGCTGGAAGCACTGGAATGCGGTACGCGGATGCTGGCGTGGCGGATTTGTTGATC 2286
Db 3482 CGGTACTGGACGCTATTGCTTCCCGCACACGCGTACTGGGCGTGGCGCTGGCATTTGATC 3541
Qy 2287 AGCCCGGCTGGCGCGGCTTGGTGGCATGAGTGGCATGAGTGGTGGCGCGCGCATCAAGCTTTA 2346
Db 3542 AACCTGGGCTGGCATCAGCAATTTGTTTATCATGGCATCGGCAAGCGTGGCTCTCGGTTTA 3601
Qy 2347 GCGGTGTTTCACTCAATGGAGCAGCATCTGCAACAGCTGCTGACGACGATCGTTACGCGC 2406
Db 3602 CTACAGGCAATGCGCTGGCGCGGAGATTCGATCGCTGCTGACTGAACAGGATACCGCG 3661
Qy 2407 TACGATGTACGATTCAGGCGCAGCTGACGCGCGCAGGCGGTTGGCAGCGTGGCGCGC 2466
Db 3662 AGCGTATGACAAAAATTCAGGCGCATTTGGGCTGCGGAGGGGACACAGCGCGCGCGC 3721
Qy 2467 ACATGCTGAGCAGGCTGTGCGCAGCAGCATGCTGCTGGCGAGGGGACCTGATGCG 2526
Db 3722 ATATTTGTTGAACAGCGCATGCG- -GACCTGTTCAGCGAGTACTCAGTGGGCGGATTTATGCA 3780
Qy 2527 CACGCAATACGATGTGATTTTGGTGGTGGCTGAGCTGGCGAATGCTGATTTGGCTGCG 2586
Db 3781 ACCGCACTATGATCTCATTTCTGGTGGTGGCGGCTGCTGGCTAATGGCTTTATCGCGCTCG 3840

Qy 2587 TCTGGGTCAATTTGAGCGCACAACTGAATGCTGTGCTGAGAGCGGATGCGCATTCGCGC 2646
Db 3841 GCTTCAGCAACAGCATCCGGATATCGGATCTTGTCTTATGAGGCGGCTCTGAGCGCG 3900
Qy 2647 AGGCAATCATACCTGGTGTTCATCAGAGGATCTCAGCGCGCAACATTCGCTGGCT 2706
Db 3901 AGGGAACCATACCTGGTCTCTTTACGAAGAGATTTAAAGCTGAATCAGCATCGCTGGAT 3960
Qy 2707 GCAACCGCTGATTTACCGTGGCTTGTAGGTTATCAGGTGGGTTTCTCGCTCGCGCG 2766
Db 3961 AGCGCGCTTGTGGTCCATCACTGGCCGACTACAGGTTCTGTTTCCCAACGCGCTCG 4020
Qy 2767 CAATCGAACGGGATTTATGTTCCATCGCATCAGGCGATTTTGGCCCGCATCTTTACGC 2826
Db 4021 CCATGTGAACAGTGGCTACTACTGCTGACCTCCCGCATTTTCGCGGGATATCTCGGCA 4080
Qy 2827 GGCGATGGTGCAGCATCTGTGAGCAAAACACAGCGGTACACAGGTAAACCCACGAGT 2886
Db 4081 ACAGTTTGGACAAATTTATGGTGCATACCGCGTTTCAGCGCTTCATGCTGAATCGGT 4140
Qy 2887 GACGTGGCGGATGGCCGTGAACCTTGTGCGCAAGTGTGATGATGCTGCGCGCTGCA 2946
Db 4141 CCAGTTAGCGATGGCCGGATTTATTCAGCCAGTACAGTATCGACGAGCGGGTTACAC 4200
Qy 2947 GCGAGCGCAATCTGAGCTGGGTATCAGGTGTTCTTGGACAAGAGTGGCAGTGGC 3006
Db 4201 GCTGATTTCTGCACTACGCTAGGATTCAGGCAATTTATCGGTCAAGAGTGGCAACTGAG 4260
Qy 3007 GCAGCGCACGCGCTGACGAGCGCATCTGATGGATGCCACCGTGCATCAGCAAGCGG 3066
Db 4261 GCGCGCATGGTTATCGTCACCGATTTATGATGGATGCGACGCTGCATCAGCAAAATGG 4320
Qy 3067 TTATCGTTTGTCTACAGCTGCGCGCTCAGCGCGCATCGGCTATTTGATGAAGATACCCA 3126
Db 4321 CTACGCTTGTGTTTATACCTTCCGCTTTCGCAACCGCATCTGCTGATCGAAGACACA 4380
Qy 3127 TTACGTTAACAGCGCGCTGGCGGAAACACCGCTCTGTCAGCAATTCGCGCATATGC 3186
Db 4381 CTACATTTGACAAGGCTATCTTCAGGCGCAACGCGCGCTCAGAACATTCGCGATATGC 4440
Qy 3187 CAATCAGCAAGCTGGAGCGCTGAGTACGCTCTGCTGAAGAGCACGCATATACCGAT 3246
Db 4441 TGCGCGACAGGTTGGCGTTACAGACGTTCTGCGGGAAGAACAGGGTGCATTCGCCAT 4500
Qy 3247 TACCTGAGCGGCAACATCGATTCGATCTGCGCAACAGCAGCGCGCCCAAGCTGACGCG 3306
Db 4501 TACGTTAACGGCGATATCTGTCAGTTTGGCAACAGCAAC- - -CGCAAGCTGTAGCG 4557
Qy 3307 CCGCGCGCGGCTGTTTCATGCAACACCGGTTACTCTTGGCGTCCGCGTGGCGCT 3366
Db 4558 ATTACGCGCGGCTGTTTCATCCGCAACCGGCTACTCCCTACCGCTCGCGTGGCGCT 4617
Qy 3367 AGCGAGTTGTAGCAGCGCTTTGCCCAACCGATGCCCTCAGCTCAGCGCAACATATCGA 3426
Db 4618 GGCGATCGTCTACGCGCTGGATGTTTACCTCTTCTCTGTGTACCAAGCATTCG 4677
Qy 3427 ACCTTTTCCCGCTCAGCAGTGGCGCAACAGCGATTTTTCGCTCTGCTTAAACCGCATGCT 3486
Db 4678 TCATTTTCCCGCAACGTTGGCAGCAACAGGGTTTTCGCGATGCTGATCGATGTT 4737
Qy 3487 GTTTTGGCGGTAAAGCGCAGCAGCGCTGGCGGTGATGAACGTTTTTACCGCTCGA 3546
Db 4738 GTTTTGGCGGACCGCGCGAGTCAAGTGGCGTGTGATGACGCTTCTATGGCTTACC 4797
Qy 3547 TGCCGGTTATAGCGCTTTTACGCGCGCACTGCGCTGCGCATTAACCGCGAT 3606
Db 4798 CGAGATTTGATGCGCGCTTTTATGCGGGAACCTCACCGTACCGCTGCGCTACCGAT 4857
Qy 3607 TCTGCGGCAAGCGCGCTGCCCATCGGTGAAGCGCTGCGCGCGCTGTTGAAATCTGCT 3666
Db 4858 TCTGAGCGGCAAGCGCGCTTCCCGTTTCCGCTTTCGCGCATTTGAGGCAATATGACGACTCA 4917

QY 3667 CGAACCGGAGAAATAATGAACGCACTTATGTTGTCGCGCAGCTTTGGCGGCTG 3726
Db |||||
QY 4918 TCGTTGAAGCGCACTACATGAACCAACTACGGTAATTGTTGGGGCTTTTGGTGGCTG 4977
Db |||||
QY 3727 GCGTTCGGCAATTCGCTCTCAAGCGCGGCGCATACCAACACCTTACTCGAGCAGCGCAC 3786
Db |||||
QY 4978 GCACTGGCAATTCGTTTACAGCGCGCAGTATCTCTGTTTGTCTTGAGCAGCGCGAC 5037
Db |||||
QY 3787 AAACCGGCGGAGCGGCTTATGTTTGAAGACAGTGGCTTTACCTTCGATGCGGACCC 3846
Db |||||
QY 5038 AAGCGGCTGGCGGCTTATGTTTATCAGGAGCAGGCGCTTTACTTTTGTATGCAAGGCGCT 5097
Db |||||
QY 3847 ACGGTGATCACCGATCCGAGCGCATCGAAGATTGTTACGCTGGCAGGAAATCGCTC 3906
Db |||||
QY 5098 ACCGTTATACCGATCCGAGCGGATGAGAACTGTTTGTCTTGGCGGTAAACAGCTT 5157
Db |||||
QY 3907 AGCGATTACGTGAGCTGATCGCGTAACGCGCTTCTATCGCTGTGCTGGGAAGATGGC 3966
Db |||||
QY 5158 AAGGATTACGTGAGCTGTGCGGTCACGCGCTTATCGCTGTGCTGGAGTCCGCG 5217
Db |||||
QY 3967 AAACAGCTTGATTACGACAAATATCAGCGGCTGCTGGAGCAGCAGATCGCCACGTTCAAT 4026
Db |||||
QY 5218 AAGGCTTCAATATACGATAACACAGCGCCAGTTAGAGCGCAGATACAGCAGTTTAA 5277
Db |||||
QY 4027 CCGCAAGATGTAGAAGGCTATCGTCAATTTCTTGCTTATTCACGTGAAGTATTTAGAG 4086
Db |||||
QY 5278 CCGCGGATGTTGCGGTTATCGAGGTTCTTGACTATTCGCGTCCGTTTCAATGAG 5337
Db |||||
QY 4087 GGTATCTGAATCGGCACTCGGCGGTTCTGAGAGTGGTGATGATGCTGGCGTCCG 4146
Db |||||
QY 5338 GGTATCTGAAGCTCGGCACTGCGCTTTTATCGTTCAAGACATGCTTCGGGCGCG 5397
Db |||||
QY 4147 CCGCAGTTGGAGCTGCAACATGCGGCGGCTACAGCATGCTGGCGAAATTTAT 4206
Db |||||
QY 5398 CCCAGTTGGCAAGCTTGAGGATGCGGCGGTTACAGTAAGTTGCGGCTACAT 5457
Db |||||
QY 4207 CAGGAGCATCATCTGCTGAGCGGTTTCTTCCATCATCTGCTGCTGGGCGTAACTCT 4266
Db |||||
QY 5458 GAGGATGACATCTTCGGCAGCGGTTTCTTTTCACTCGCTCTTAGTGGGGGGAATCG 5517
Db |||||
QY 4267 TTTGCAAGCTCATGATCTATACCTTAATCATGCTGGAGCTGAAATGGGCGGTGG 4326
Db |||||
QY 5518 TTTGCAACCTCGTCCATTTATACGCTGATTCAAGCTTGAACCGGAATGGGCGCTGG 5577
Db |||||
QY 4327 TTTTCGCGCGGCGCACCGCGCTGTCAGGCGATGCGCGCATGCTTTCAGGACTTG 4386
Db |||||
QY 5578 TTTTCACGCGGTGGAACCGGTGGCTGGTCAATGGCATGATCAAGCTGTTTCAGGATCTG 5637
Db |||||
QY 4387 GCGGCGAGCTGTTACTGAATGCGGAAGTGAGCAGCTGGAAACGAGCGGCAATCGCAT 4446
Db |||||
QY 5638 GCGGCGGAAGTCGTGCTTAAACGCGCGGTCACTCATATGGAACCGTTGGGACAAAGATT 5697
Db |||||
QY 4447 AGCGGCTTCAGTTAGGCGGACACGCTTCGATGCGCGCTGTCGCTCCCAATGCC 4506
Db |||||
QY 5698 CAGGCGGTGAGTGTGAAGACGCGACAGCTTTGAACCTGCGCGTGGCTGCGACGCT 5757
Db |||||
QY 4507 GAGCTGCTGATACCTACGACAACTGCTTCGCCACCATCCGCTGGCAATGAACGTTGG 4566
Db |||||
QY 5758 GATGTTGTACATACCTATCGGATCTGCTGTCTCAGCATCCCGCAGCGCTAAGCAGGCG 5817
Db |||||
QY 4567 ACATCGGTGAAGCTGAAGCGCATGAGCACTCGCTGTTGTACTCTATTTTGGCTGAT 4626
Db |||||
QY 5818 AAAAACTGCAATCCCAAGCGTATGAGTAACTCACTGTTGTACTCTATTTGCTCTCAAC 5877
Db |||||
QY 4627 CAGCGCATGAACGCTCGGCGACACACCGCTGTTTGGCGCGGTTATCGTGAAGTTG 4686
Db |||||
QY 5878 CATCATCAAGTCACTCGCCCATCATACGCTGTTTGGGCGCACGCTACCGTGAATG 5937
Db |||||
QY 4687 ATCGATGAGATTTTCAACAGCAGCAGCTGGCAGACGATTTTTCATTTACCTGCAAGCG 4746
Db |||||
QY 5938 ATTCAGGAATTTTAAACCATGATGCTGCTGAGGATTTTTCGCTTTATTTTACACGCA 5997
Db |||||
QY 4747 CCCTGCAGCAGCATCGTCTGCTGGCACCGCCCGCTGCGGCACTTTTATGTTAGCG 4806

Db 5998 CCTTGTTGTCACGGATCCGTCACCTGCGACCGGAAGGTGCGGCACTTATGTTGCTGGCG 6057
QY |||||
Db 4807 CCGTTCGCGCATCTCGGCAACCGCTGACATCGCTGCAACAGGAAGAGACCGCGCTTGGCG 4866
QY |||||
Db 6058 CTTGTTCCACACTTAGCGACCGGCAACCTGCACTGGCGGTAGAGAGACCCGACCTGGCG 6117
QY |||||
Db 4867 GATCGAATTTTTCGCTTATCTGAGCAGCACTACATGCTCGGGAATTACGTCAGCAATTAGTG 4926
Db |||||
Db 6118 GATCGTATTTTTCGCTTACCTTACCTTGAACAATTTACATGCTGGCTTGGAGCCAGTTGGTG 6177
QY |||||
QY 4927 ACACACAGATTTTACCGCGCTTGTATTTTCGCGACACGCTGATGCTGATGCTGCTGCG 4986
Db |||||
Db 6178 ACGCACGATTTTACCGCGCTTGTATTTCCGCGACGAGCTCAATGCTGCGCAAGGTTTCG 6237
QY |||||
QY 4987 GCGTTTTGCTGAGCGGATTTTGAACGMAAGCGCTGTTCCGCGCGCATATAACCGCGAT 5046
Db |||||
Db 6238 GCTTCTCGGTTGAACCTATTTGACCCAGAGCGCTGTTGTTCCGACACATATAACCGGAT 6297
QY |||||
QY 5047 GCGATATACGAATCTCTATCTGTTGGTGGTTCGCTGACGATCCAGGCGGGGCTGCC 5106
Db |||||
Db 6298 AAGCACATTTGATATCTTATCTGTTGGCGCAGGCAACCATCTGCGCGGCGCATTTCC 6357
QY |||||
QY 5107 GCGTGTATCGTTGCGCAAGGCGCACCGCAGGCTGATGCTGGAGGATCGCGCGAATGA 5166
Db |||||
Db 6358 GCGGTAATCGGCTCGGCGAAGGCGACGCGAGGCTTAAATGCTGGAGGACCTGATTTGACGA 6417
QY |||||
QY 5167 ATCGACAGCTTTTACTTTGAGCAAGTAAACGCAAACTTACGCGGCTGGGCTCGAAGATTTTCG 5226
Db |||||
Db 6418 AT---AGCTATTTACTGAATCATGCTCGAACCCTGCGGTTGGCTCGAAGAGCTTTG 6474
QY |||||
QY 5227 CCACCGCGCAAGCTGTTTGTATGACGAGCGCGCGCAGCAGCTGATGCTGATGCTG 5286
Db |||||
Db 6475 CGACTGATCGACGCTTTTTCGACGCGCAAAACCCGCTCGCAGCGCTGCTGATGCTTTACGAT 6534
QY |||||
QY 5287 GGTGCTGCTCATGCGATGATGATTGATGCGGCAAAACGCTGGGCGAAGGCGGACGCGAGC 5346
Db |||||
Db 6535 GGTGCGCGCACTGCGAGCAGCTCATTTGACGATCAAAACACTTGGGCTTTTCATGCGCGACG 6594
QY |||||
QY 5347 ATGCGCTCGAAGACGCGCAGCAGCATATGAGCATCTGCAAAATTTGAAACCGCGCGCT 5406
Db |||||
Db 6595 CCTCTTCGAGATGCTTCGAGCAGCGCTGACGAGCTTGAATGAATGAACCGCTCAGGCT 6654
QY |||||
QY 5407 ACAGGCGCGCACATGATGAACCGGCGTTTGGGCGTTTACGAAAGTGGCGATCATTC 5466
Db |||||
Db 6655 ACGCGGTTTCGAAATGACAGAGCCCGCTTTGCGCGCTTTTCAGAGGTTCGCGATGGCG 6714
QY |||||
QY 5467 ACCAGCTGCGCGCAACAACTGCGGCTTGTATCATCTGGAAGGCTTCGCTATGATGCAAGCA 5526
Db |||||
Db 6715 ATGATATCGCTCCGCGCTTACCGTTTCGACCATCTGGAAGGTTTTCGCTATGATGCGCG 6774
QY |||||
QY 5527 ACGAATATTACGCGAGCTTTCGATGACAGCTGCGTTACTGCTATCATGCTCGGCGCTGG 5586
Db |||||
Db 6775 AAACGCGCTACTGACACTGAGCAGTACGCTGCGTTATTTGCTATCATGCTCGCGCTGTTG 6834
QY |||||
QY 5587 TCGGTTTGTATGATGATGCGCGCTTAATGCGCGTGCAGCAGGAGCGGTGCTCGATCACGCT 5646
Db |||||
Db 6835 TGGGCTGATGATGCGCAAAATTTATGGCGGTTTCGCGATTAACCGCAGCTCGATCGCGCT 6894
QY |||||
QY 5647 GCGATTTAGGACTGCGGTTTCCAGCTCACTAACTTGGCGCGCAATTTGTAAGATGCGG 5706
Db |||||
Db 6895 GCGATCTCGGCTGCTTTTCCAGTTTGAACAATGCGCGTGAATTTGCTGACGATGCTC 6954
QY |||||
QY 5707 AAAATGCTGCTGCTATCTGCGCAATCTGCTGCTGATGCGCGGATTAACGCGCGATA 5766
Db |||||
Db 6955 AGGTGGCGCGCTGTTATCTGCTGAAAGCTGCTGGAAGAGAGGACTGAGCAAGGCA 7014
QY |||||
QY 5767 CGCTGACTGCAACGCAACATGCTGCGCGCTCGCTCACTGCGAGCGGCTTTTGTGGCGG 5826
Db |||||
Db 7015 ATTATGCTGCGCAGAAACCGCAGGCTTTAAGCGGATTCGCGGCGACTGTTACGGG 7074
QY |||||
QY 5827 AGCGGGAACCTTATTCATCTGCGGCGCATTCGCGTTTACCGGTTTACCGCTGCGCTCGG 5886

Db 7075 AAGCGAACCTTATTACGTATCATCAATGCGCGTCTGCGACAAATTACCTTACGCTCGG 7134
Qy 5887 CQTGGCCATCGTACGCTCGCGGCTTATCGCGAATATGCGGTCGCAAGTTACGACG 5946
Db 7135 CTTGGGCCATCGCGACAGCGAAGACGAGTACCGTAAATTTGGCGTGAAGTTGAACAGG 7194
Qy 5947 CCGGTGTGACGCGCTGGGATTCACGCGACGACGACAGTAAAGTGAAGAACTGGCGCTGC 6006
Db 7195 CCGGTAAAGCGCTGGGATCATCGCCAGTCAACGTCACCGCGGAAATTAACGCTTT 7254
Qy 6007 TGGTAAAGGGGAGGTTTGGCGATCACTTCGCGTGTCTCGTCTCGAAGCGCGTCCGG 6066
Db 7255 TGTGACGCGCATCGGTACGAGGATTTACCTCCGATGAAGACGATTCACCCCGCTCG 7314
Qy 6067 CTGGTCTGTGGAGCGTCTCGTTGATTTAGTCTCGTGAACGCTGGCGGAGCGTGGCTTG 6126
Db 7315 CTCATCTCTGGAGCGCC-----GATCTAGCGCGATGCCCTTCTCAGCGTGGCGCTG 7368
Qy 6127 CAGCTTATTTCAGCGTGGCGCTAGAGGAACCAAGACGACGCGCTTACGCGCGG 6186
Db 7369 AAGTTTAGATTAACGGTGGCGGTACAGAAACCAAGGACACGCGCCCTCTTTCCCT 7428
Qy 6187 CACCGCATGATGCGGTGGCCCATGTATTAAGCGCTTAAAGATAGCCCTTTCGCGGGAT 6246
Db 7429 TACAGCATGATGATACGCTGGGCCATGTATAACCGTTTTCAGTAGCCTTTCGCGGTAT 7488
Qy 6247 ATAGCGGAACGCGAGCTGTGATGACACGAGCGCATGTCGACACATGAAGTAGAGCGCGC 6306
Db 7489 GTAGCGGAACGCGAGCGCTGTGTACCACTGCTGTGGACCAATAAATACAGTAAACC 7548
Qy 6307 GTACGTGCTCATTCCGCGACCAATCACTGAGCGGCCACATGCTTGCACACGACATA 6366
Db 7549 ATAAGCGTTCATCGCTGCAACCAATCCACTGGAGCGGCCAGATTCCTGTACTCCGA 7608
Qy 6367 AATCAGCACATCGCCAGTACCGCAAAACCAACCGCATAAAGATCGTTGAGCTCAAACTT 6426
Db 7609 ATCAGGGCAATCGACACATGGCGNATACCGCATAGAGATCGTTAACTTCAATGC 7668
Qy 6427 ACCGTGTGCGTTTCATGTGCGACAGATGCCAGCGCCCATCCCAACCGTGCATGATGA 6486
Db 7669 GCCTTTACCGGTTTCATGATGTGAAGATGCCAGCGCCCAACCCAGCGCTGCATGATGA 7728
Qy 6487 TTTATCGACAGCGCGCTACGATTCATCCATCCACCGCTTGCACCAAGATAGCAC 6546
Db 7729 TTTATGTGCGAGTACGAAACCACTTCCATGCGGACCGGTGACAAACACGATCAGGCG 7788
Qy 6547 GTTCCATACAGAGCAATTTGCTGCCA 6574
Db 7789 ATTCCAAATCCACACATATTTCTCAA 7816

RESULT 4

US-10-734-778-40
; Sequence 40, Application US/10734778
; Publication No. US20040209365A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Suh, Wonchul
; TITLE OF INVENTION: PARALLEL CHROMOSOMAL STACKING OF TRAITS IN BACTERIA
; FILE REFERENCE: CL2026 US NA
; CURRENT APPLICATION NUMBER: US/10/734,778
; PRIOR FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/434773
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 8609
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pPCB15

US-10-734-778-40
Query Match 37.4%; Score 2619.6; DB 18; Length 8609;
Best Local Similarity 64.7%; Pred. No. 0;
Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;
Qy 373 CAGCTGCGACCAATTTTACAGCGCATCTTGAACATTTACTGCTGCCGACGACGAAAGC 432
Db 1627 CAGTTGCTGGCTGATATCGATAGCGCCTTGATCATGTTACTGCCGTTCAGGGTAGCGG 1686
Qy 433 GATCCGCTGCGTCCGCGATGCGTCCGGAAACGCTGGCGCAGGCAAAAGCTATTTCCTCT 492
Db 1687 GATTGTGTGGTGGCGCGATGCGTGAAGCACGCTGGCACCGGCAAAAGCTATTTCGTCG 1746
Qy 493 TTATTACTGCTGCTGGCAGCGCGCATATGGGTTCGAGCTGACGCAAAATGGCGTTCTC 552
Db 1747 ATGCTGCTGTTATTAAACAGCGCGCATCTTGGCTGTGCGATCAGTCAACGCGGATTTAC 1806
Qy 553 GATCTCGCTGTGCGATGCGAATGCTGCACGCGCATCGCTGATTTCTGGATGACATTTCCC 612
Db 1807 GATTTAGCTTGGCGGTTGAAATGTCATGCTGCTCGCTGATTTCTGGATGATATGCC 1866
Qy 613 TCGATGGATTAACCGCAGATGCTGCTGCTCGCCCTACCGTCATCGCGAATTTGGTGAA 672
Db 1867 TGCATGGACGATGCGCAGATGCGTGGGGCGTCCACCATTTCAACGACGATACGGTGA 1926
Qy 673 AACGGCGATTTCTCGCGCCCATCGCTGTGTAGCCGCGCATTTGAAAGTGAATGCCATT 732
Db 1927 CATGTGGCGATTTCTGGCGCGTCTGCTTACTCAGCAAAAGCTTTGGGGTGTATGCGGAG 1986
Qy 733 GCACCGGTTTGCCTGCCATACATTAATCTGAAGCGATTTGCTGAACCTCTCCGCTGCCGTC 792
Db 1987 GCTGAAGTCTGACCGCCATAGCCAAACTCGCGCGGTGTGCGAGCTGTCTCACTCGGAT 2046
Qy 793 GSCCTGCGAGGCTTAGTCAAGGGCAATTTCCAGGATCTGACGACGCGCACGACGAGCCGC 852
Db 2047 GGCATGCGAGGCTGCTGTTTTCAGGCGCGATTTAGGACCTCTCGAAGCGCATAAACCCCG 2106
Qy 853 AGCCCGAAGCGATGCCATGACCAACGAATCTGAACAAACGAGCGTCTGTTTCGCGCACG 912
Db 2107 AGCGCGGATGCCATATCTGCTAACCAATCAGTTTAAACCCAGCACGCTGTTTGGCGGTCA 2166
Qy 913 CTGCAATGGCGGCGATTTGCCGCTCACGCTTACCGAGGTGCGGCAAGACCTTAGCTTC 972
Db 2167 AGCAATATGGCGTTCATTTGCGGCCAACCGCTCTCGAAGCGCTGAGAACCTTGATCGT 2226
Qy 973 TTTGCGCCAGGATTTTGGCGCAGCGCTTTCAACTGCTCGACGCTCGCGACGCTTTCGCA 1032
Db 2227 TTTCTGCTCGATCTCGGCCAGGCGCTTTTCACTGCTTGAAGCTTACCGATGCGATGACC 2286
Qy 1033 CACACCGTAAAGATGTGCAACAGATCAGGGCAATCCACGCTGGTACAGATGCTCGGT 1092
Db 2287 GATACCGGCAAGACATCAATCAGGATGCAAGTAAATCAACGCTGCTCAATTTATTAGGC 2346
Qy 1093 GCTGACGCGCGGAACGCTGCGCTGCGCATCACTGCGCAGCGCAGATGACACCTTGC 1152
Db 2347 TCAGCGCGGTCGAAAGAACGCGCTGCGCAGCATTTTGGCGCTGCGCAGTGAACACCTTTC 2406
Qy 1153 TCGCGCTGCCATCGCGGATCGCCACTCGCCAA-----TATATCAACGCGCTGTTTAAATCAA 1209
Db 2407 GCGGATGCGCAAAACGCGCATTTCCACACCCAACTTTTATTTACGCGCTGGTTGACAA 2466
Qy 1210 CAGCTAGCGATATTCAACTGAGCGCGGCTCAGCGGTCAGCGGCACTTTTGGCGTATGCGCG 1269
Db 2467 AAACCTCGCTGCGCTCAGTTAAGGATGCTGCA-----TGAGCCATTTTGGCGTATGCGCAC 2521
Qy 1270 CGCGCTCTACAGCCACTTTTACGCGTTGCGAGCGGTTAGCAAAACGCTGCTGGCGCGG 1329
Db 2522 CGCCCTTTTTCAGCCATGTTTTCGCGCTGCAAAACCTTCTCTCAGGAATTTAGTGGCGCGG 2581
Qy 1330 GGCATCGCATCACTTTCATCCAGGAACGCGATGCGCGCACTTTTCTTAGCGACGACGCA 1389
Db 2582 GTCACTGCTTACGTTTTTTTTCAGCAACATGACTGCAAGCGCTGTTAACGCGGACGATA 2641

QY 1390 TCGATTTTGTTCGGCTCGCCCAACAGACGATCTCTCGCGTTCTCGCGCCCGCTGTTC 1449
 Db 2642 TCGGATTCAGACCGTCGACATGCAACAGCATCTCTCGCGTTCTCTATCGCACCTGTTC 2701
 QY 1450 ATCGCTCGCTCGCGCGCGCTGTCTCGTTTTCGGTGTATCGAGATCTCGCGTCT 1509
 Db 2702 ACCTCGCGCGCACCACTCGACCCCTCGATGTTTACGACTGATCAATGAATGSCAGTA 2761
 QY 1510 GCACCGATATGTGTGCGCGCAATGCTCTCGGTA CTGAAGCATTTGAACATCGATGGCG 1569
 Db 2762 CCACCGATATGTTTTCGCGGAATGCTCGCGCTCTTTCATGCGTTGAGATAGAGGCG 2821
 QY 1570 TGATCGCGACGAATGGAAGCGCGCGGAGTGTCTGCTGAAGCGCTGCATCTGCGT 1629
 Db 2822 TGATCGTTGATCAATGGAAGCGCGGAGTGTCTGCTGAGAGCGTCAAGTCTGCGCT 2881
 QY 1630 TTGTTTTCGCTGCGCTTTCGCGTCAATTCGTGAAGCGCGGATTCGCTTTCGCTGA 1689
 Db 2882 TTGTTTTCGCTGCGCTTTCGCGTCAATTCGTGAAGCGCGGATTCGCTTTCGCTGA 2941
 QY 1690 TGCCCTTCGCTTTCGACAGGATGACAAAGCGCTGAAGCTTTTCAGGCGCAGCAGATA 1749
 Db 2942 TGCCCTTCGATACGACGATGCGCTTCGCGAAGCTATACCAACGAGCGAATAA 3001
 QY 1750 TCTATGATCGCATCATGCTGCTGCTCAGCGAGCTGATCTCAACACAGCGCGGCTTTA 1809
 Db 3002 TTATGACTGCTGATGCGACGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 3061
 QY 1810 ATTGACGAGCGCGCGGATTAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1869
 Db 3062 GTTACCGCGCTGGAATAACTGCTATCTGTTTCTTCTCACTGCGCAAAATCAGCCAGT 3121
 QY 1870 TGCTGCGCGCTTGTATTTTCAGCTGACAACTGCGCGCTGCTGCTGCTGCTGCTGCTG 1929
 Db 3122 TGATCCCGCAACTGGAATTTTCCCGCAAGCGCTGCACTGCTTTCATGCGGTGAC 3181
 QY 1930 CACTTCGCG - - - CCGCGTTCCTCTGCGCGCTCCATGCGCGCTGCGCGCGCTGCTG 1986
 Db 3182 CGTTACGGCAACCCGAGGAGCGCGGCTCATCAACTTCTTATTTTCGCTCCCGGACA 3241
 QY 1987 AGCGGTGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2046
 Db 3242 AACCCCGTATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3301
 QY 2047 ATCTGCGCAGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2106
 Db 3302 CCATTCGCAAGCTTGGAAAGGTTGATGCGCAGTTACTGTTGGCAGCTGTTGGCGGC 3361
 QY 2107 TAAACGCGCAACAGACGATCAGCTGAGCTGCTGCGCGCGCTGCTGCGAGATTTG 2166
 Db 3362 TCTCAGCCACGAGGAGGTAATGSCCGCGCGGAGACATTCAGGTTGTGATTTG 3421
 QY 2167 TCGATAGCGCGGAGCCCTACAGCAGCGAGCTGTTTATCACTCATGCGGTTAAACA 2226
 Db 3422 CCGATCAATCCGAGCACTTTTCAAGGACAGTTGAAATCAACATGTTGGATGAATA 3481
 QY 2227 GCGCGTGAAGCACTGGAATGCGTACGCGGATGCTGCGCTGCTGCGGATTTGCTGATC 2286
 Db 3482 CGGTACTGAGCGTTATGCTTCCGACACAGCTACTGCGCTGCGCTGCGCTGCGCTG 3541
 QY 2287 AGCCCGCGTGGCGCGCATGAGTGGATGAGCTGCTGCGCGCGCTGCGCGCATCAGCTT 2346
 Db 3542 AACCTGCGGTGGCATCAAGATTTTATCATGCGATCGGCAAGCGTGTCTCGGTTA 3601
 QY 2347 GCGGTGTTCACTGAGCAGCATCTGCAACAGCTGCTGACGAGCATGCTTAAGCGC 2406
 Db 3602 CTAACAGCATTCGCTGCGCGGAGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3661
 QY 2407 TACGATGCTCAGCGATTCAGCGCAGCTGACGCGCAGCGGTTGCGGCTGCGCGC 2466
 Db 3662 AGCGTATGACAAATTCAGGCGCATTTGCTGCTGCGAGGCGGACACAGCGCGCGC 3721

QY 2467 ACATCGTCAGCAGCGCTGTCTCCAGCAAGTCTGTCTGCGGAGGCGGAGGCGGCTGATGCG 2526
 Db 3722 ATATTTGTTGAACAGGCGATGCG - GACCTGTACGCGAGTACTCAGTGGGAGGATTTATGCA 3780
 QY 2527 CAGCGAATACGATGATTTTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2586
 Db 3781 ACCGCACTATGATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3840
 QY 2587 TCTGCGTCAATTCAGCCCAACTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2646
 Db 3841 GCTTCAGCAACAGCATCCGATATGCGATCTTGTCTTATGAGGCGGCTCTGAGCGCG 3900
 QY 2647 AGCAATCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2706
 Db 3901 AGGAACCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3960
 QY 2707 GCAACCGCTGATTAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2766
 Db 3961 AGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4020
 QY 2767 CAATCTGAGCGGCGATTTATTTCCATTCGATCAGCGCATTTTTCGCGCGCATTTTACGC 2826
 Db 4021 CCATGTGAACAGTGGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4080
 QY 2827 GCGATGCTGACGATCTGTGGAACAACACAGCGCTGCTGCTGCTGCTGCTGCTGCTG 2886
 Db 4081 ACAGTTTGGACAACATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140
 QY 2887 GAGCTCGCGGATGCGCGTGAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2946
 Db 4141 CCAGTTAGCGGATGCGCGGATTTATTCATGCGAGTACAGTATGCTGCTGCTGCTGCT 4200
 QY 2947 GCCGACCCACATCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3006
 Db 4201 GCCTGATCTGCACTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4260
 QY 3007 GCAGCGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3066
 Db 4261 CGCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4320
 QY 3067 TTATCGTTTGTCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3126
 Db 4321 CTACCGCTTGTGTTATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4380
 QY 3127 TTACGTTAACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3186
 Db 4381 CTACATTGCAAGGCTAATCTTCAGGCGGAAAGCGGCGGCTGCTGCTGCTGCTGCTGCTG 4440
 QY 3187 CAATCAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3246
 Db 4441 TCGCGGACAGGCTTGGCGCTTACAGCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4500
 QY 3247 TACCTGAGCGCGCAACATTCGATTCGATTCGCAACAGCAGCGCGCGGCGGCTGCTG 3306
 Db 4501 TACGTTAACGCGCGGATATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4557
 QY 3307 CTGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3366
 Db 4558 APTACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4617
 QY 3367 AGCGGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3426
 Db 4618 GSCCGATGCTCTCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4677
 QY 3427 AGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3486
 Db 4678 TCACCTTTGCGGAGCAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4737
 QY 3487 GTTTTGGCGGTAAGCGCGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3546
 Db 4738 GTTTTGGCGGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4797
 QY 3547 TGCGCGGTTAAATGCGCGCTTTTACGCGGCGCACTGCGCTGCTGCTGCTGCTGCTGCTG 3606

Db 4798 CGAGGATTTGATGCGCGCTTTTATGCGGAAACTCACCGTGACCGGATCGGCTACGCAT 4857
Qy TCTGTGCGGCAAGCGCGCTGCCATCGGTGAAGCGCTGCGCGCTGTTGAAATCTGT 3666
Db TCTGAGCGGCAAGCGCGCTTCCCGTTTTCGCGCATTTGCAGGCAATTTATGACGACTCA 4917
Qy CGAACCGAGGAAAGAAATGAAACGCACTTATGTGATTTGGCGGAGGCTTTGGCGGCTG 3726
Db TCGTTGAAGAGCGACTACATGAACCAACTACGCTAAATGTGTGCGGCTTTGCGGCTG 4977
Qy GCGCTGGCAATTCGCTGCAAGCGCGGCATACCAACCACTTACTCGAGCAGCGCGAC 3786
Db GCACTGGCAATTCGTTACAGCGCGCAGGTATTCCTGTTTGTGCTGTGAGCAGCGCGAC 5037
Qy AAACCGGGCGGACGCGCTTATGTGTTGAGGACAGTGGCTTTACCTTCGATCCGAGCCC 3846
Db AAGCGGGTGGCGGCTTATGTGTTATCAGGACGAGGCTTTTACTTTTGTGATCGAGCCCT 5097
Qy ACGGTGATCACCGATCCGAGCGGCATCGAAGAGTGTTCACGCTGGCAGGAAATCGCTC 3906
Db ACCGTTATCACCGATCCGAGCGGATTTGAAGAACTGTTTGTCTCTGCGCGGTAAACAGCTT 5157
Qy AGCGATTACGTCGAGCTGATCGGTAAAGCCCTTCTATCGCTGTGCTGGGAAGATGSC 3966
Db AAGGATCTGAGCTGTGCGGTGCGGTACCGCTTATCGCGTTTATCGGCTGTGCTGGAGTCCGCG 5217
Qy AAACAGCTTTGATTACGCAATTAATCAGCGCGCTGTGGAGCAGCAGATCGCCAGTTCAT 4026
Db AAGGCTCTCAATACGATAACAGCAGCGCCAGTTAGAGCGCAGATACAGCAGTTTAA 5277
Qy CCGCAAGATGTAGAGGCTATCGTCAATTTCTTGTGCTATTCAGTGAAGTATTTAGAGAG 4086
Db CCGCGCATGTTGCGGGTATCGAGCGTTTCTTGACTATTCGCTGCGGTATTTCAATGAG 5337
Qy GGTATCTGAACTCGGACGCTGCGTTTCTGAGGTCGCGTGCATGCTGCGCTCGCG 4146
Db GGTATCTGAACTCGGACGCTGCGTTTCTGAGGTCGCGTGCATGCTGCGCTCGCG 5397
Qy CCGCAGTTGGGACGCTGCAAGCATGGCGCAGCGTCTACAGCATGGTGGCGAAATTTAT 4206
Db CCCAGTTGGCAAGCTGCAGGATGGCGCAGCGTTTACAGTAAAGTTGCGCGCTACAT 5457
Qy CAGGAGATCATCTGGTTCAGCGGTTTCTCTTCATCTGCTGCTGGGCGGTAACTCT 4266
Db GAGGATGAGCATCTTCGCGCAGCGTTTCTCTTCTCTGCTCTTATGAGGAGGAGTTCG 5517
Qy TTTGCAAGTCTCATCTATACCTTAAATTCATGCGCTGGAGCGGTGAATGGGCGGTGCG 4326
Db TTTGCAAGTCTCATCTATACCTTAAATTCATGCGCTGGAGCGGTGAATGGGCGGTGCG 5577
Qy TTTCCGCGCGGCGCACCGCGCGCTGTGTGAGGCGATGGCGGCTGTTCGAGGACTTG 4386
Db TTTCCAGCGGGTGAACCGTGGCTGTGTCAATGCGATGATCAAGCTGTTCAGGATCTG 5637
Qy GCGCGGAGCTGTACTGATGCGGAAAGTGAAGCAGCTGGAAACCGAGCGCAATCGCAT 4446
Db GCGCGCAAGTGTGCTTAAACCGCGGTGAGTCAATGAAACCGGTGGGGGCAAGATT 5697
Qy AGCGGCTTCAGTTAGGCGGCGACGACGCTTCGATGCGCGCTGCTGCTGCTCCATGCC 4506
Db AGCGGCTTCAGTTAGGCGGCGACGACGCTTCGATGCGCGCTGCTGCTGCTGCTGCTGCT 5757
Qy GACGTGGTGCATACCTACGACAACTGCTTCGCCACCATCCGCTGGCAATGAAGTGGG 4566
Db GATGTTGTACATACCTATCGCGATCTGCTGTCTCAGCATCCCGCAGCGCTAAGCAGGG 5817
Qy ACATCGCTGAAGGTTAAGCGATAGCAACTCGCTGTTTGTACTCTATTTTGGCTGAT 4626
Db AAAAATCTGCAATCCAAAGCGTATGAGTAACTCACTGTTTGTACTCTATTTTGGTCTCAAC 5877
Qy CAGCGCATGAACGCTCGGACACACCGCTCTGTTTGGCGCGGTATTCGCTGAGTTG 4686

Db 5878 CATCATCAGATCAACTCGCCCATCATACCGTCTGTTTTGGGCCACGCTACCGTGAACGTG 5937
Qy ATCGATGAGATTTTCAACAGCAGCAGCTGGCAGACGATTTTTCATCTTTACCTGACGCG 4746
Db ATTCAGGAAATTTTAAACCATGATGCTGCTGAGGATTTTTCGCTTTTATTTACAGCA 5997
Qy CCCTGACAGGAGTCCGCTGCTGCGACCGCCCGCTGGGAGCTTTTATGTTGTAGCG 4806
Db CTTGTGTGTCGAGATCCGCTGCTGCGACCGGAAGGTTGGGAGCTATTTATGTTGTGGCG 6057
Qy CCGGTGCGCATCTCGGACACGCTGACATCGACTGCGCAACAGAAAGGACCGGCTTGGCG 4866
Db CTTGTTCACACTTAGGCAAGCGGAACTCGACTGGCGGTAGAGGACCCGACTGCGCG 6117
Qy GATCGAATTTTGTCTTATCTGAGCAGCACTACATGCTCCGGGATTCGTGAGCAATTTAGTG 4926
Db GATCGAATTTTGTCTTATCTGAGCAACATTTACATGCTCTGGCTTGGCAAGCAGTTGGTG 6177
Qy ACACACAGATGTTTACCGCGTTTGTATTTTCGCGACACGCTGCAATGCCCTGCAAGGCTCG 4986
Db ACACACAGATGTTTACCGCGTTTGTATTTTCGCGACACGCTGCAATGCCCTGCAAGGCTCG 6237
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5046
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5297
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5106
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6357
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5166
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6417
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5226
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6474
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5286
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6534
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5346
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6594
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5406
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6654
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5466
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6714
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5526
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6774
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5586
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6834
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5646
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6894
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5706
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 6954
Qy GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 5766
Db GCGTTTTCGCTGGAGCGGATTTGACGCAAGAGCGCTGCTCCGCGCATATAACCGCGAT 7014

```
QY 5767 CGCTGACTGCACCGCAACATCGTCAGCGCTCGCTCACTGCGACGCGGTTTGTGGCGG 5826
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7015 ATTATGCTGGCGCAGAAACCGCAGGCTTAAGCGGTATCGCCGGCGACTGTACGGG 7074
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5827 AGCGGAACCCATTATCACTCGGCGCGATCGGTTTACCGGTTTACCGTGCCTCGG 5886
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7075 AAGCGGAACCCATTATCACTCGGCGCGATCGGTTTACCGGTTTACCGTGCCTCGG 7134
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5887 CGTGGGCGCATCGCTACGCTCGCGCGTTTACGGAATTTGGCGTCAAGTTACAGCAG 5946
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7135 CCGTGGCCATCGCGACGAGAGGTTACCGTAAATTTGGCGTGAAGTTGAACAGG 7194
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 5947 CCGGTGTGACCGCTGGGATTCACGGCAGCGCACAGTAAGAGTGAATAAATCTGGCGCTGC 6006
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7195 CCGGTAAAGAGGCGAGTTTGGGATCACTTGGCGTGTCTCGTCTCTGAAACGCGTCCGG 6066
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7255 TGCTGACGCGATCCGCTCAGGCGATTTACTTCCCGATGAAGACGATATCCACCCCGCTCG 7314
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6067 CTGCTGTGGCAGCGTCTCGTTGATTTACGTCGCGAGCTGGCGCAGCGTGGCTTG 6126
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7315 CTCATCTCTGGCAGCGCC-----GATCAGCGCATGCCCTTCTCAGCGTCCGCTG 7368
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6127 CAGCTTATTCAGCGGTGGCGGTAGAGAAACCAACGACGACGCGCTTCAACGCCGCG 6186
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7369 AAGTTTAGATAACGGTGGCGGTACAGAAACCAAGGACACGACGCGCTCTTTCCCT 7428
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6187 CACCGATGATCGATCGGTGGCGCATATGATTAAGCGCTTAAGATAGCTTTGGCGGGAT 6246
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7429 TACAGCATGATCATACGCTGGGCGCATATTAACCGTTTCAAGGTAGCCTTTGCGCGGTAT 7488
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6247 ATAGCGGAACCGCGCTGATGACACGAGCGCATCTGCACCATGAAGTAGAGCGCGCC 6306
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7489 GTAGCGGAACCGCGCGGTGTGTACCGTCCGTCGTGGACCATAAATACAGTAACC 7548
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6307 GTACGCTGTCTATTCGGCACCAATCCACTGACAGCGGCCACATGCTTGCACACGACATA 6366
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7549 ATAAGCGGTCTATGCTGCACCAATCCACTGAGCGGCCAGATTCCTGACTGCCGAAGTA 7608
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6367 ATACGACAAATCGCCAGTACCGCAACACCGCATGAAGATCGTTGAGCTCAAACTT 6426
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7609 AATCAGGGCAATTCGACAAATGGCGAATACCGCATAGATCGTTAACTTCAATGCG 7668
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6427 ACCGCTGTGCGTTTATGTGGGACAGATGCCAGCGCCCATCCCAACCGTGCATGATGA 6486
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7669 GCCTTTACGGGTTTATGATGTGAAGATGCCAGCCCAACCCAGCGGTGCATGATGA 7728
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6487 TTTATGCGACAGCGCGTACGATTTCCATCCACCGACCGGTTGCCAACAGATAGCNC 6546
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7729 TTTATGTGCGAGTGCAGCAACCACTTCCATGCGGACCAAGTGCACAAACACGATCAGGCG 7788
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 6547 GTTCCATACCGAGCATGTTGCTCCA 6574
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 7789 ATTCCAAATCCAAACATTAATTTCTCAA 7816
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 5

```
US-10-734-936-37
; Sequence 37, Application US/10734936
; Publication No. US20040209370A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Co., Inc.
; APPLICANT: Suh, Wonchul
; TITLE OF INVENTION: Method for Chromosomal Engineering
; FILE REFERENCE: CL1878 US NA
; CURRENT APPLICATION NUMBER: US/10/734,936
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/434602
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
```

; SEQ ID NO 37

; LENGTH: 8609

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: Plasmid pPCB15

US-10-734-936-37

Query Match 37.4%; Score 2619.6; DB 18; Length 8609;

Best Local Similarity 64.7%; Pred. No. 0;

Matches 4015; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;

QY 373 CAGCTGACGCAAAATTTACAGCGCATCTTGAACATTTACTGCTGCGGACACCAAGC 432

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1627 CAGTTGCTGGCTGATATCGATAGCCGCTTGTATCAGTTACTGCGCGTTTCAGGGTGAOCGG 1686

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 433 GATCGCTGGCTGCGCGATGCGTGGCGGAACGCTGGCGAGGCAACGATTTTCGTCCT 492

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1687 GATTGTGGGTGCGCGATGCGTGAAGGACGCTGGCACCGGGCAACGATTTTCGTCGG 1746

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 493 TTATTACTGCTGTCGCGAGCGCGCATATGGTTTGCAGCTGACGCAAAATTTGCGTTC 552

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1747 ATGCTGCTGTTATTAAACAGCGCGCATCTTGGCTGTGCGATCAGTCACGGGGATTA 1806

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 553 GATCTCGCTGTGACGTGGAATGCTGCACGCGCATGCTGATTTCTGGATGACATTC 612

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1807 GATTAGCTGCGCGTTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1866

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 613 TCGATGATTAACGCGCAGATGCGTGGTTCGCTTACCGCTACCGTCATCGCAATTTGGTGAA 672

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1867 TGCATGACGATGCGCAGATGCGTGGGGGCTCCACCATTCACACGACGATCGGTGAA 1926

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 673 AACGTGCGATTTTCGCGCGCATCGCGTGTAGCGCGCATTTTGAAGTGAATTCGCTATT 732

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1927 CATGTGCGATTTTCGCGCGCGTCTTACTAGCAAAAGCGTTTGGGGTGAATTCGCGAG 1986

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 733 GCACCGGTTTTCGCTGCCATATAAATCTGAAGCATTTGTAACCTTCGCTGCGCTG 792

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1987 GCTGAAGTCTGACGCGCGATAGCCAAACTCGCGCGGTGTCGAGCTGTCACCTGCGATT 2046

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 793 GGCCTGACGGGCTTAGTGCAAGGGCAATTCAGGATCTGACGACGCGCACGACGACGCGC 852

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2047 GGCATGACGGGCTTGGTTTCAGGGCCAGTTTAAAGACCTTCGGAAGGCGATAAACCCGCGC 2106

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 853 AGCCCGGAAGCGATCGCATGACCAACGAATGAAACACGCGTGTCTGTCGCGCCACG 912

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2107 AGCGCGGATGCCATCTGCTAAACCAATCAGTTTAAACACGACGCTGTTTGGCGGTCA 2166

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 913 CTGCAAAATGGCGGATTTGCGCTGACGCTTTCACGCGAGTGGCGCAAGACTTAGCTTC 972

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2167 ACGCAATGGCGTCCATTTGCGCCCAACGCGTCTTCGGAAGCGGTGAGAACCTGCATCGT 2226

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 973 TTGCGCCAGGATTTGGCGCAGCGGTTTCAACTGCTGACGACCTGCGCGACGCGGTTGCAAA 1032

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2227 TTCTCGCTCGATCTCGCGCAGGCGCTTTCAGTTGCTGACGATCTTACCGATGGCATGACC 2286

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1033 CACACCGTAAGATGTGCAACAGGATCAGGCAATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2287 GATACCGCAAGACATCAATCAGATGACGATAATCAACGCTGCTCAATTTATTAGGC 2346

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1093 GCTGACCGCGCGGAACGTCGCTGCGCATCACCTGCGCGACGCGACAGATGCAACCTTCC 1152

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2347 TCAGCGCGGTGCAAGACGCTGCGACAGCATTTGCGCTGCGCGAGTGAACACCTTTCC 2406

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1153 TGCCTGCTGCGGATCGCGGATCGCACTCGCAAA---TATATGACGCGCTGTTTAAATCAA 1209

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2407 GCGCATGCCAAAACGGCCATTTCACACCACTTTTATTCAGGCTGTTTACACAAA 2466

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1210 CAGCTAGGATATTCAACTGAGCGCGGCTCAGCGGCTGGCCACTTTTGGGTGATCGCGC 1269

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2467 AAATCGCTGCGGTGAGTTAAGATGCTGCA-----TGAGCCATTTTGGGTGATCGCAC 2521

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1270 CGCGCTCTACAGCCACTTTTCAACGCTTTCAGCGGTTGACGGGCTTAGCACAACGCTGTCGGCGCGC 1329

DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 2522 CCCCCTTTTACGCAATGTTCCGGCTCTGCAAAAACCTTGCTCAGGAATTAGTGGCCGCG 2581
Qy 1330 GCCATCGCATCATTTATCCAGCAAGCCGATGCCCGCACTTTGCTTAGCGACGAACGCA 1389
Db 2582 GTCATCGTTAGTACGTTTTCAGCAACATGACTGCAAAAGCGCTGGTAAGCGGCAGCGATA 2641
Qy 1390 TCGATTTGTTGCGCTGCGCCAAAGACGATCTGCGGTTGCGTGGCGCCCGTGTGTC 1449
Db 2642 TCGGATTCAGAACCGTCGGAATGCAAAACGATCTCCCGGTTCTTATGCGACCTCTGCTGC 2701
Qy 1450 ATCGGCTGCGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509
Db 2702 ACTGCGCGCGCACCACTCGGACCCCTCATGTTACGACTGATCAATGAATGGCACGTA 2761
Qy 1510 GCACCGATATGCTGTCGCGCAACTGCCCTGCGGTACTGAAAGCATTTGAACATCGATGGCG 1569
Db 2762 CCACCGATATGCTTTGCGCGGAATGCCCGCGCTTTTCATGCGTTGCGATAGAGGCG 2821
Qy 1570 TGATCGCGACGAATGGAAGCGCGGCGGAGTGGTCTGCTGAAGCGCTGCATCTGCGGT 1629
Db 2822 TGATCGTTGATCAAAATGGAAGCGCGGAGTGCAGTAGTCGCAAGAGCGTCAGGCTCTGCCGT 2881
Qy 1630 TTGTTTCCGTTGCGCTGCGCTTGGCGCTCAATCGTGAAGCGGATTCGCTTGGGTGA 1689
Db 2882 TTGTTTCCGTTGCGCTGCGCTTGGCGCTCAACCGCGAAGCGGTTTGGCTCTGCGGTGA 2941
Qy 1690 TGCCCTTCCGTTTGCACAGGATGCAAAAGCGCTGAAAGCGTTTTCAGGCGCAGCAGCGATA 1749
Db 2942 TGCCCTTCGAGTACGGCACCGGATCGGCTCGGGAAGCTATACCAACAGCGAAGAAA 3001
Qy 1750 TGTATGATCGATCATGCTGCTGACGGCGAGTGAATCTCAAAACAGCGCGGGGTTTA 1809
Db 3002 TTTATGACTGGTGTGACGCTGCAATCGTGTGATCGCGCATCATGCAATCAGAAATGG 3061
Qy 1810 ATTTGACGAGGCGCGGATTTACATCAGTGCTGTCGCGCTGCGCAAGCAATCAGCCAGA 1869
Db 3062 GTTTAGCCCCGCTGAAATGCAATCATTTGTTTCTCCACTGGCAAAATCAGCCAGT 3121
Qy 1870 TGTGCGCGGCTTTGATTTTCAACGCTCAGCAACTGCGCCCTGCTATCACGCGTGGGCG 1929
Db 3122 TGATCCCCGAATGATTTTCCCGCAAGCGCTGCCAGCTGCTTATGCGGTGGAC 3181
Qy 1930 CACTCCGCG --- CCGCGTTTCTCTGCGCGCTCATGCGCGCTTGGCGCGCTGCGCTC 1986
Db 3182 CGTTACGGCAACCCAGGCGGACCGCGGCTCATCAACTTCTTATTTTCGTTCCCGGACA 3241
Qy 1987 AGCGGTTGTTTATGCTCGCTGGGTACGCTGCAAGGCCATCGCTTCCGCTGTTTCTGC 2046
Db 3242 AAACCCGCTATTTTGCCTCGCTGGGACCCCTCAGGGACATCGTTATGGCCTGTTTCAGGA 3301
Qy 2047 ATCTGCGCGAGCGCTGCCCGCAGCTGCGGCTATCGCTGATCGCCCATTTGCGGGAT 2106
Db 3302 CCATCGCCAAAGCTTGGGAAGGTGGATCGCAGTTACTGTTGGCACTATGTCGCGGCC 3361
Qy 2107 TAAACCGCAACAGACGATCAGCTGGAGCTCGCTGGCGCGCTGGGTGACGGAATTCG 2166
Db 3362 TCTAGCCACAGCGGAGGTGAATGCGCGCGGCGGACATTCAGGTTGTTGGATTTTG 3421
Qy 2167 TCGATCAGCGCGCAGCCCTACAGCACGCGCAGCTGTTTATCACTCATGCGCGGTTAAACA 2226
Db 3422 CCGATCAATCCCGACGACTTTCACAGGCACAGTTGACATCAATCATGTTGGGATGAATA 3481
Qy 2227 GCGCGCTGGAAGCACTGGAATGCGGTACGCCGATGCTGGCGCTGCGGATTTGTTGATC 2286
Db 3482 CCGTACTGACCGCTATTTCTCCGCAACCGCTACTGGCGCTGCGCTGGCATTTGATC 3541
Qy 2287 AGCCGCGGTGCGCGCGCATTTGATGGCATGACGTTGGTTCGCCCGCATCACGCTTTA 2346
Db 3542 AACCTGGCGTGCATCACGAATTTGTTTATCATGGCATCGCAAGCGTGGCTCTCGGTTTA 3601
Qy 2347 GCGGTGTTTCATCAATGAGGACGATCTGCAACAGCTGCTGACCGACGATCGTTAGCGCG 2406

Db 3602 CTACAGCCATGCGCTGCGCGCGAGATTCGATCGCTGCTGACTAAACACCGATTACCCGC 3661
Qy 2407 TACGATGTCAGCGATTCAGCGCGAGCTGCAAGCGCGCAGGCGGTTGCCAGCGTGC CGCG 2466
Db 3662 AGCGTATGACAAAAATTCAGGCGCATTTGCGTCTGCGAGGCGGCACACAGCGCGCGCG 3721
Qy 2467 ACATCGTCAGCAGCGCGCTGTGCCAGCAGCAAGTGTGCTGGCGGAGGGACCTGATGCG 2526
Db 3722 ATATTGTTGAAACAGCGCGATGCG - GACCTGT CAGCCAGTACTCAGTGGGCGAGGATTATGCA 3780
Qy 2527 CACGCAATACGATGATGATTTTGGTGGTGGTCTGCAATGCGGATGCTTGAATTCGCTGG 2586
Db 3781 ACCGACATATGATCTCATTTCTGGTGGTGGCGGTCTGCTAATGCGCTTATGCGCTCG 3840
Qy 2587 TCTGCGTCAATTTGACAGCCAACTGAAATGCTGTTGCTGGAGAGCGATGCGCATTCGCGC 2646
Db 3841 GCTTCAGCAACAGCATCCGATATGCGGATCTTGTCTTATGAGGCGGCTCTCTGAGCGCG 3900
Qy 2647 AGGCATCATACCTGGTGTTCATCAGCGATCTCAGCGCGCAACAACTTCGCTGGCT 2706
Db 3901 AGGGAACCATACCTGGTCTCTTTCAGAAAGAGGATTTAAACGCTGAATCAGCATCGCTGGAT 3960
Qy 2707 GCAACCGCTGATTAACGCTGGTGGTCAAGTTCAGGTTGCTGCTGCTGCTGCTGCTGCTG 2766
Db 3961 AGCGCGCTGTGGTCTCATCTGCGCCGACATCAGGTTGCTTCCCGCAACGCGCTCG 4020
Qy 2767 CAATCTGACGGGATTAATGTTCCATCGCATCAGCGGATTTTGGCCGCACTCTTTACGC 2826
Db 4021 CCATGTGAACAGTGGCTACTACTGCTGACCTCCCGCATTTTCGCGGGATACTCCGGCA 4080
Qy 2827 GCGATGGGTGACGATCTGTGGAACAAACAGCGCTACAAAGGTAAACACAGCGAT 2886
Db 4081 ACAGTTTGGACAACATTTATGGCTGATACCGCGGTTTTCAGCGCTTCATGCTGAATCGGT 4140
Qy 2887 GACGCTGCGGATGSCCGTGAACTTTGCTGCGCAAGTGTGATGTTGTCGCGGCTGCA 2946
Db 4141 CCAGTTAGCGATGGCCGATTTATTCAGCCAGTACAGTATGAGGAGCGGGTTACAC 4200
Qy 2947 GCCAGCGCACATCTGACGCTGGGTTATCAGTGTGTTCTTGGAAGAAGTGGCAGTGGC 3006
Db 4201 GCCTGATTTCTGCACTACGCTAGGATTCAGGCAATTTATCGGTCAAGGATGGCAACTGAG 4260
Qy 3007 GCAGCGCAGCGCTGACGAGCGGATCCTGATGATGCGCACCGCTGATCAGCAAGCGCG 3066
Db 4261 CGCGCGCATGTTTATCGTCAACGATTCATGGAATGCGACGCTGATCAGCAAAATGG 4320
Qy 3067 TTATCGTTTGTCTACAGCTGCGCTCAGCGCGCATCGGCTATTGATTTGAAGATACCCA 3126
Db 4321 CTACGCTTTGTTTATACCTGCGCTTTCCGCAACCGCATCTGCTGATGCAAGACACAC 4380
Qy 3127 TTACGTTAACAGCGCGCTGGCGGAGAACACCGCTCTGTCAGCACATGCGCGACTATGC 3186
Db 4381 CTACATTCAGCAAGGCTAATCTTCAGGCGAAGCGGCGCTCAGAACATTCGCGATTATGC 4440
Qy 3187 CAATCAGCAAGCTGGAGCTGATGATCGCTGCTGGTGAAGGACGCGCATTTACCGAT 3246
Db 4441 TCGCGCAGCAGGTTGGCGGTTTACAGACGTTGCTGCGGGAAGAACAGGGTGCAATGCCAT 4500
Qy 3247 TACCTGAGCGCAACATCGATTCGATTCGCAACAGCAGCGCGCGCAAGCTGCGCGG 3306
Db 4501 TACGTTAACGGCGGATTAATCGTTCAGTTTGGCAACAGCAAC --- CGCAAGCTGTAGCGG 4557
Qy 3307 CTTGCGCGCGGCTGTTTTCATGCGCACACCGGTTTACTCTTTCGCGCTCCGCGTGGCGGT 3366
Db 4558 ATTACGCGCGGCTGTTTTCATCGCAACCGGCTTACTCTTCCCTACCGCTGCGGTTGGCGCT 4617
Qy 3367 AGCGAGTTGGTAGCGGCTGTGCCCAACCGATGCCCTCAGGCTCAGCAACATATCGA 3426
Db 4618 GCGCGATGCTCTCAGCGCGCTGGATGTTGTTTACCTCTTCTCTGTTTCCAGCAGATTCG 4677
Qy 3427 ACCTTTTCCGCTCAGCAGTGGCGGCAACAGGATTTTTCGCTGCTGCTAAACCGGATGCT 3486
Db 4678 TCATTTTCCCAAGCAACGTTGGAGCAACAGGGGTTTTTCCGCATGCTGATTCGATGTT 4737

QY	3487	GTGTTTGGCCGTAAGCCGACGACGCTGCGCGGTGANGCAACGTTTTHACCGGCTCGA	3546	QY	4567	ACATCGCTGAAGCGTAAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCCCTGAAT	4626
Db	4738	GTGTTTGGCCGACCGCGGAGTCAAGCTGCGGTGATGATGAGCGGTTTCTATGCTTACC	4797	Db	5818	AAAAACTGCATCCAAAGCGTATGAGTAATCACTGTTTGTACTCTATTTTGGTCTCAAC	5877
QY	3547	TGCCGGTTAATTAGCCGCTTTTAGCCGGGCACTGCGCTGCGGATAAACGCGGAT	3606	QY	4627	CHGCCCATGAACAGCTCGCGGCAACACCGCTGTTTGGCCCGGTTTATCGTGAATTG	4686
Db	4798	CGAGGATTTGATTTGCCGCTTTTATGCGGGGAAACTCACCGTGACCGATCGGCTACGAT	4857	Db	5878	CATCATCAGATCAACTCGCCCATCATACCGCTGTTTGGGCCACGCTACCGTGAACATG	5937
QY	3607	TCGTGCGGCAAGCCGCGCTGCCCATCGGTGAAGCGCTGCGCGCTGTTTGAATTTCTGT	3666	QY	4687	ATCGATGAGATTTTCAACAGCAGCAGCTGGAGACGATTTTTCATCTTTACCTGACACGG	4746
Db	4858	TCGTAGCGGCAAGCCGCGCTTCCGTTTTCGCGCAATTGCGGCAATTAATGACGACTCA	4917	Db	5938	ATTACAGAAATTTTAAACCATGATGCTTGCGCTGAGGATTTTTCGCTTTATTACACGCA	5997
QY	3667	CGAACGAGGAAAGAAAATGAACGCACTTATGATGATGGCGCAGGCTTTGGCGGCTG	3726	QY	4747	CCCTGACAGCAGCATCCGCTGCTGGCACCCCGCGCTGCGGACGCTTTTATGTTAGCG	4806
Db	4918	TCGTGAAGAGCGACTACATGAACCAACTACGCTAAATGTTGGTGGGGCTTTGGTGGCCTG	4977	Db	5998	CTTTGCTCAAGATCCGTCACCTGCGCACCGGAAGGCTGCGGACGCTATTATGTTGCTGGC	6057
QY	3727	GGCTGGGATTCGGCTGCAAGCGCGGCAATACCAACACCTTACTCGAGCAGCGCGAC	3786	QY	4807	CCGGTGGCGATCTCGGCACCGCTGACATCGATCGCAACAGGAAGACCGGCTTGGCG	4866
Db	4978	GCACTGGCAATTCGTTTACAGCCCGGAGTATTCCTGTTTGTCTTGAGCAGCGCGAC	5037	Db	6058	CCTGTTCCACACTTAGGCAAGCGGAACTCTGATGCGCGGTAGAAGAACCCCGACTGGC	6117
QY	3787	AAACCGGGCGGACGCGCTTATGTTTGAAGACAGTGGCTTTTACCTTCGATGCGCGACCC	3846	QY	4867	GATCGAATTTTCTTATCTGAGAGCAGCACTACATGCGCGGATTAACGTCAGCAATTAGTG	4926
Db	5038	AAGCCGGTGGCGGCTTATGTTTATCAGGACGAGGCTTTTACTTTTGTATGTCAGGCGCT	5097	Db	6118	GATCGATTTTGTACTACCTTGAGCAACATTAATGCTGGCTTGGGAAGCCAGTTGGTG	6177
QY	3847	AGGTGATCACCGATCCGAGCGGCATCGAGAGTGTTCACGCTGCGAGGAAATCGCTC	3906	QY	4927	ACACACAGAAATGTTTACGCCGTTTGTATTTTCGCGACACGCTGATGCCCATCAGGCTCG	4986
Db	5098	ACCGTTATCACCGATCCGAGCGGATTTGAAGAACTGTTTGTCTGCGCGGTAACAGCTT	5157	Db	6178	ACGCACTGATGTTTACGCCGTTTCTGATTTCCGCGACGAGCTCAATGCTTGGCAAGGTTTCG	6237
QY	3907	AGCGATTACGTCAGCTGATGCGGTAAAGCCCTTCTATFCGCTGTGCTGGGAAGATGCG	3966	QY	4987	GCCTTTTCGCTGGAGCGGATTTTGAAGCAAGCCCTGTTCCGCGCGCATAAACCGGAT	5046
Db	5158	AAGGATTTACGTCAGCTGTGCGGTACGCGCTTATTCGCTGTGCTGGGAGTCCGCG	5217	Db	6238	GCCTTCTCGGTTGAACTTATCTGACCCAGAGCGCTTGGTTCGACCAACATTAACCGGAT	6297
QY	3967	AAACAGCTTGAATACGACAAATATCAGCCGCTGCTGGACGACGATCGCCAGCTTCAAT	4026	QY	5047	GCCGATATCAGCAATCTTATCTGTTGGTGGTCCGCTGATGATCCAGGCGCGGGGTGCC	5106
Db	5218	AAGGTTTCAATTAACGATTAACGACGAGGCGGCTTTAGAAAGCAGATACAGCAGTTTAT	5277	Db	6298	AAGCACATGATTAATCTTTATCTGTTGGCGACGACCCATCTCTGGCGGGCAATTTCC	6357
QY	4027	CCGCAAGATGTAAGAGCTATCGTCAATTTCTTGCTTATTCAGCTGAAGTATTTAGAGAG	4086	QY	5107	GGCTGATCGGTTGCGCAAGGCGCACCGCGCTGATGCTGAGAGATCGCGCGAATGA	5166
Db	5278	CCGCGGATGTTGCGGGTATTCAGAGGTTCTTGTGACTATTCGCTGCGGTATTCATGAG	5337	Db	6358	GGCGTAATCGGCTCGGGAAGGCGACGCGAGGCTTAATGCTGAGAGACCTGATTGACGA	6417
QY	4087	GGTTATCTGAACCTCGGCAAGTGGCGTTCGTCAGGTCGCTGACATGCTGCGCTGCGG	4146	QY	5167	ATCGACAGCTTTACTTTGAGCAAGTAACGCAAAACCATGCGGCTGGGTCGAAAGTTTCG	5226
Db	5338	GGCTATCTGAAGCTCGGCACTGTGCTTTTATCGTTCAAAAGACATGCTTGGGCGCGG	5397	Db	6418	AT---ACGTCAATTAATGATCATCCGCTCGAAACCATGCGGTTGGCTCGAAAGCTTTG	6474
QY	4147	CCGAGTTGGGAGCTGTGCAAGATGCGCAGCGCTCTACAGCATGCTGCGGAAATTTAT	4206	QY	5227	CCACCGCGCCAAAGCTGTTGATGCCAGCGCGCCCGCAGCAGCTGATGCTGATCGCT	5286
Db	5398	CCCAGTTGGCAAGCTGCAAGGATGGCGAGCGCTTTTACAGTAAGTTGCGGCTACAT	5457	Db	6475	CGATGCTACGACGCTTTTCGACGCCAAACCCCGCTCGCAGCGCTGATGCTTTTACGAT	6534
QY	4207	CAGGACGATCATCTGCGTCAGGCGTTCCTTCCACTCATTTGCTGGTGGCGGTAACTCT	4266	QY	5287	GGTGTGCTCACTGCGATGATGATGATGGGCAAAAGCGCTGGGCGAAAGGCGGCGCAGC	5346
Db	5458	GAGGATGAGCATCTTCGCGAGCGGCTTTCTTTTCACTGCTCTTTAGTGGGGGGAATCCG	5517	Db	6535	GGTGGCGGCACTGCGACGACGCTCATTTGACGATCAACACTGGGCTTTTCATGCCGACCG	6594
QY	4267	TTTGCAACGTCATGATCTATACCTTAATTCATGCGCTGAGGCGTGAATGGGCGGTGCG	4326	QY	5347	ATGCGCTCGAAGACGCGCAGGCACTATGACGATCTGCAAAATTTGAAACCCCGCGCGCT	5406
Db	5518	TTTGCAACCTCGTCCATTTATACGCTGATTCACGCGTTAGAACGGGAATGGGCGGTCTG	5577	Db	6595	CCTCTTCGACATGCTTGCAGCGCGCTGACGAGCTTGAATGAAACCGCTCAGGCGCT	6654
QY	4327	TTTCCGCGCGGCAACCGCGCTGCTGCGAGGCAATGGCGGACTGTTTCGAGGACTTGG	4386	QY	5407	ACAGCGCGCGCATATGATGAAACCGGCTTTAGGCGGTTTTCAGGAAGTGGCGGATCTTC	5466
Db	5578	TTTCCACGCGGTGAACCGGTGCGCTGCTCAATGGCATGATCAAGCTGTTTCAGGATCTG	5637	Db	6655	ACGCGGTTGCGAAATGACGAGCGCGCTTTTGGCGGTTTCAGGAGTTCGCGATGGCGC	6714
QY	4387	GGCGCGGAGCTTTTACTGAATGCGGAAGTGAAGCGAGCTGGAACACCGAGCGCAATTCGAT	4446	QY	5467	ACCAGCTCGCGCAACAACTGGCGTTTGTATCATCTGAAAGGCTTCGCTATGATGACCGCA	5526
Db	5638	GGCGGCAAGTCTGCTTAAACCGCCGGTCAAGTCAATATGGAACCCGTTGGGGCAAGAT	5697	Db	6715	ATGATATCGCTCCCGCTACGCTTCGACCATCTGGAAGGTTTGGCCATGATGTTGCGCG	6774
QY	4447	AGCGCGTTTCAATTAGAGGCGGACGAGCTTCGATGCGCGCTGTTGGCTTCCAATGCC	4506	QY	5527	ACGAACTATTCGCGAGCTTCGATGACGCTGCGCTTACTGCTATCACCTCGCGGCGGTGG	5586
Db	5698	CAGGCGGTGCAATTGGAAGACGCGCAGCGGTTTGAACCTGCGCGGTGGCGTTCGACGCT	5757	Db	6775	AAACGCGCTTACCTGACACTGGAACGATACGCTGCGTTTGTCTATCACCTCGCGGCTGTG	6834
QY	4507	GACGTGGTGATACCTACGACAAACTGCTTCGCCACCATCCGCTGGCAATGAACCGTGGC	4566	QY	5587	TCGCTTTGATGATGCGCGCGTAAATGGCGCTGCGCGAGCAAGACGCGTGTCTCGATCAACG	5646
Db	5758	GATGTTGTACATACCTATCGGATCTGCTGCTCTCAGCATCCCGCAGCGGCTAAGCAGCGG	5817	Db	6835	TGGGCTGATGATGCGGCAAAATTTATGGGCTTCGCGATTAACGCCACGCTCGATCGCGCT	6894
				QY	5647	GCGATTTAGGACTGCGGTTCCAGTCACTAAATTTGGCGGCGACATTTGTAGAGATGCGG	5706


```
Db 6895 GCGATCTCGGCTGCTTCCAGTTGACCAATTCGCGGTGATATTGTGACGATGCTC 6954
Qy 5707 AAAATGGTCTGCTTATCGCGCAATCTGTGCTCGATCAGCGCGGATTAAGCGCGGATA 5766
Db 6955 AGGTGGCGGCTGTTATCTGCTGAAAGTGGCTGGAAGAGAAAGACTGACRAAGCGA 7014
Qy 5767 CGCTGACTGACCGCAACATGCTGAGCGCTCGCTCACTGCGCAGCGGTTTGTGCGGG 5826
Db 7015 ATTATGCTGGCGCAGAAAACCGCGAGGCTTAAAGCCGTATCGCCGGGCGACTGGTACGGG 7074
Qy 5827 AGGGGGAACCTTATTATCACTCGGCGCATCCGGTTTACCGGTTTACCGCTGCGCTCGG 5886
Db 7075 AAGCGGAACCTTATTACGTATCATCATGCGCGGTCTGGCAATACCTTTACGCTCGG 7134
Qy 5887 CGTGGCCATCGCTACGCTCGCGCTGTTTATCGCGAAATTTGGCGTCAAAAGTTTCAGCAG 5946
Db 7135 CCGTGGCCATCGCGACGGAAGCAGGTGTACCGTAAATTTGGCGTGAAGTTGAACAGG 7194
Qy 5947 CCGTGTGACCGCTGGGATTCACGGCAGCGACCAAGTAAGGTGAAAACTGGCGCTGC 6006
Db 7195 CCGGTAAAGCGGCTGGGATCATCGCCAGTCCACGTCACCGCCGAAAAATTAACGCTT 7254
Qy 6007 TGGTGAAGGGGACGTTTGGCGATCACTTCGCGGTGTCTCGTCTGAAACCGGCTCGG 6066
Db 7255 TGCTACGCGATCCGCTCAGGCGATTACTTCCCGATGAAGACGTATCCACCCGCTCGT 7314
Qy 6067 CTGCTGTGGCAGCGCTCTCGTTGATTTTACGCTGACGCTGGCGCAGCGTGGCTG 6126
Db 7315 CTCTCTCTGGCAGCGCC-----GATCTAGCGCGATGCCCTTCTCTCAGGCTCGCTG 7368
Qy 6127 CAGCTTATTCAGCGTGGCGGTAGAGAAACCAAAACGACAGCGCTTACGCGCGCG 6186
Db 7369 AAGTTTAGATAACCGTGGCGGTACAGAAAAACAAAGGACACGCGAGCCCTCTTTCCGCT 7428
Qy 6187 CACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6246
Db 7429 TACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7489
Qy 6247 ATAGCGGAACGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6306
Db 7489 GTAGCGGAACGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7548
Qy 6307 GTACGTGCTATTCGCGCACCAATCCACTGAGCGGCGCACATGCTTGGACACCGACATA 6366
Db 7549 ATAAGCGGTGATGCTGACCAATCCACTGAGCGGCGCACATGCTTGGTACTGCGCAAGTA 7608
Qy 6367 AATCAGCACATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6426
Db 7609 AATCAGGGAATCGACACAATGGCGAATACCAACGCGATGATGATGATGATGATGATGAT 7668
Qy 6427 ACCGCTGTGCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6486
Db 7669 GCCTTATCGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7728
Qy 6487 TTTATGCGACAGCGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6546
Db 7729 TTTATGTCAGTGCAGCAACCACTTCCATGCGGACCAAGTGAACACGATCAGGCG 7788
Qy 6547 GTTCCATAACGAGATGTTTCGTCCA 6574
Db 7789 ATTCCAAATCCAAACATAATTTCTCAA 7816
```

RESULT 6

```
US-10-735-442-64
; Sequence 64, Application US/10735442
; Publication No. US20040219629A1
; GENERAL INFORMATION:
; APPLICANT: E. I. duPont de Nemours and Company, Inc.
; APPLICANT: Suh, Wonchul
; APPLICANT: Rouviere, Pierre
; APPLICANT: Cheng, Qiong
```

```
; TITLE OF INVENTION: Increasing Carotenoid Production in Bacteria Via Chromosomal
; FILE OF INVENTION: Integration
; FILE REFERENCE: CL2027 US NA
; CURRENT APPLICATION NUMBER: US/10735.442
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/434618
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 64
; LENGTH: 8609
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pPCB15
US-10-735-442-64
```

Query Match 37.4%; Score 2619.6; DB 18; Length 8609;

Best Local Similarity 64.7%; Pred. No. 0;

Mismatches 0; Conservative 0; Mismatches 2169; Indels 24; Gaps 7;

Qy 373 CAGCTGCAGCAAAATTTTACAGCGCATCTTGAACATTTTACTGCTGCGCGACAGCAAGC 432

Db 1627 CAGTTGCTGGCTGATATCGATAGCGCTTTCATCAGTTACTGCGGTTTCAGSGTGAGCGG 1686

Qy 433 GATCGGTGCTGCGCGATGCGTCCGGAACGCTGGCGCAGGCAAAAGTATTCGTCT 492

Db 1687 GATTGTGGTGGCGCGATGCTGAAGCAGCGCTGCGCAGCGGCAAAAGTATTCGTCT 1746

Qy 493 TTATTACTGCTGCGCAGCGCGATGATGGTTCGAGCTGACGCAAAATGGCGCTTC 552

Db 1747 ATGCTGCTGTTTAAACAGCGCGGATCTTGGCTGTGCGATCAGTCACGGGGATTTACTG 1806

Qy 553 GATCTGCGCTGTGAGTGAATGTGACGCGCATCGCTGATTTCTGGATGACATTC 612

Db 1807 GATTAGCTGCGGTTGAATGGTGCATGCTGCTGCTGATTTCTGGATGATTC 1866

Qy 613 TCGATGATATAACGCGCAGATGCGTGGTTCGCTTACCGTGCATCGCGAATTTGGTGA 672

Db 1867 TGCATGACGATGCGCAGATGCGTGGGGGGTCCACCATTTACACGACGATCGTGAA 1926

Qy 673 AACGTGGCGATTCGCGCGCATCGCTGCTTAGCCGCGATTTGAAGTATTCGCATT 732

Db 1927 CATGTGGGATTTGCGCGCGTCTTTACTTCAGAAAGCGTTTGGGGTATTC 1986

Qy 733 GCACCGGTTTGCCTGCCATACATAAAATCTGAAGCGATTGTGAACTCTCCGCTGCCGTC 792

Db 1987 GCTGAAGTCTGACCGCGATAGCCAAACTCGCGGTTGCGAGCTGCTCCACTGGATT 2046

Qy 793 GGCCTGCAAGGCTTAGTCAAGGGCAATTCAGGATCTGCACGACGCGCACGACGACGCGC 852

Db 2047 GGCATGCAAGGCTTGGTTTCAAGGACCTCTCGGAAGGCGATAAACCCTCC 2106

Qy 853 AGCCCGAAGCGATCGCCATGACCAACGAACTGAACACGCGTCTGTTTCGCGCGCG 912

Db 2107 AGCGCGATGACATCTGCTAAACCAATCAGTTTAAACACGACGCTGTTTTCGCGCTCA 2166

Qy 913 CTGCAAAATGGCGGCGATTGCGCTGACGCTTTCACGCGAGGTGCGGCAAGACTTACGTTTC 972

Db 2167 ACGCAATGGGCTCCATTGCGGCCAAACGCGTCTCGGAAGCGGTGAGAACTTGCATCGT 2226

Qy 973 TTCGCCCAAGGATTTGGCGCAGCGGTTTCACTGCTGACGACCTCGCGCAGCGGTTGCAAA 1032

Db 2227 TTTCTGCTCGATCTCGCGCAGCGCTTTTCACTGCTTGAAGTCTTACCGATGGCATGACC 2286

Qy 1033 CACACCGGTAAAGATGTGCACCGCATCAGGCAAACTCCAGCTGCTGACAGATGCTCGGT 1092

Db 2287 GATACCGCAAGACATCAATCAGATGCAAGTAAATCAACGCTGCTCAATTTATAGGC 2346

Qy 1093 GCTGACGCGCGGAAAGCTGCGCTGCGCATCACTGCGCAGCGCAGATGACACCTTGC 1152

Db 2347 TCAGGCGGTGCAAGACGCTGCGACAGCATTTGCGCGCTGGCCGACACCTTTC 2406

QY 1153 TCGGCTCCCATCGCGCATCGCCACTCGCCAA---TATATGACGCGCTGTTTATCAA 1209
 Db 2407 GCGGATGCCAAAACGGCCATTCACCAACCACCACTTTTATTCAGGCTGTTTGAACA 2466
 QY 1210 CAGCTAGCATATCAATGAGCGGGCTCAGCGGTGGCCACTTTTGGGTGATCGGC 1269
 Db 2467 AAATCGCTGCGCTCAGTTAAGGATGCTGCA-----TGAGCANTTTGGGTGATCGAC 2521
 QY 1270 CGCGCTCTACAGCCACTTTACGCGTTGCAAGGCTTAGCAAAACGCTGCTGCGCGCG 1329
 Db 2522 CGCCCTTTTCAGCCATGTCGCGCTCTGCAAAACCTTGTCTCAGGAATTAGTGGCGCG 2581
 QY 1330 GCATCGCATACATTCATCAGCAAGCGGATGCCGCACATTTGCTTAGCCACGACGCA 1389
 Db 2582 GTCATCGTTTACGTTTTCAGCAACATGACTGCAAAAGCGCTGTAACGGGACGATA 2641
 QY 1390 TCGATTGTTGCTCGGCTCGGCCAAAGACGCTATCTCGCGGTTGCTGGCGCCGCTGTTGC 1449
 Db 2642 TCGGATTCAGACCGTTCGGAATGCAACGCACTCTCCCGTTCTTATCGCACCTGCTGC 2701
 QY 1450 ATCGGCTGCGCTCGCGCGCGCTGTCGCTGTTTGGGTGATCGACGATCTCGGTCCT 1509
 Db 2702 ACTGCGCGCGCACCCACTCGGACCTCGATGTTACGACTGATCAATGAATGGCACGTA 2761
 QY 1510 GCACCATATGCTGCGCGCACTGCTCGGCTACTGAAGCAATTGAACATCGATGGCG 1569
 Db 2762 CCAGCATATGCTTTGCGCGGAATGCGCGCGCTTTTCATGCTTGCAGATAGAGGCG 2821
 QY 1570 TGATCGCCGACGAAATGGAAGCGCGCGGAGTGGTGGTGAAGCGCTGCAATCGCGGT 1629
 Db 2822 TGATCGTTGATCAATGGAAGCGCGGAGTGCAGTAGTGCAGAGCGTCAAGTCTCGCGT 2881
 QY 1630 TTGTTTCGCTGCGCTGCGCGCTGCGGTCATTCGCTGACGCGGATTCGCTTGGGTA 1689
 Db 2882 TTGTTTCGCTGCGCGCTGCGGCTCAACCGGAAACCGGTTGCTCTGGCGGTGA 2941
 QY 1690 TGCCTTTCGCTTTGACAGGATGACAAAGCGCTGAACGTTTTCAGGCGACGACGATA 1749
 Db 2942 TGCCTTTGAGTACGGCACCGGATGCGGCTCGGGAACGCTATACCAACGAGCAAAA 3001
 QY 1750 TCTATGATGCATTCGCTGCTGACGGGAGTGTATCTCAACACGCGCGGGCTTTA 1809
 Db 3002 TTTATGACTGCTGATGCGAGCTGACGATCGTGTGATCGCGATCATGATGAGAAATGG 3061
 QY 1810 ATTGAGGAGCGCGCGGATACATCAGTGCCTGTCGCGCTGGCAAAATCAGCCAGA 1869
 Db 3062 GTTTAGCCCGGTGAAACTGATCATTTGTTTCTCCACTGGCACAAATCAGCCAGT 3121
 QY 1870 TGTGCGCGCTTTGATTTTCCAGTCAGCAACTGCGCGCTGCTATCAGCGGTGGGC 1929
 Db 3122 TGATCCCGAACTGGATTTTCCCGCAAGCGCTGCGCAGACTGCTTTTCATGCGGTTGGAC 3181
 QY 1930 CACTCGCG---CCCGGTTTCTCGCGCGCTCCATGCGCCCTGCGCAGCGCTCGCTC 1986
 Db 3182 CGTTACGGCAACCCAGGGAGCGGGGTATCAACTTCTTATTTTCGTCGCCGACA 3241
 QY 1987 AGCCGCTGCTTTATGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046
 Db 3242 AACCCGTAATTTTGGCTCGCTGCGCACCTGCGAGGACATCGTTATGCGCTTTTCAGGA 3301
 QY 2047 ATCTGGCGAGGCTGCGCGCTGCGGCTATCGTGGTGTATCGCCATTTGGGGGAT 2106
 Db 3302 CCATCGCAAGCGCTGCAAGAGGTGATGCGCAGTTACTGTTGGCACATGTTGGCGGC 3361
 QY 2107 TAAACGCGAACAGACGATCAGCTGAGTCTGCTGCGCGCGGCTGGGTGACGGAATTCG 2166
 Db 3362 TCTCAGCACGCGAGGAGTGAATGCGCCGCGGCGGGACATTCAGGTTGTGGATTTTG 3421
 QY 2167 TCGATCAGCGCGAGCCCTACAGCACGCGGAGCTGTTTATCATCTACGCGGTTTAAACA 2226
 Db 3422 CCGATCAATCGCAGCACTTTACAGGCGACAGTTTGAACATCACATGTTGGGATGATA 3481
 QY 2227 GCGCGCTGGAAGCACTGGAATGCGGTACCGCGTCTGCGGCTGCCGATGCTTTTGTATC 2286

Db 3482 CGGTACTGACGCTATTTGCTTCCGCACACCTACTTGGCGCTGCGCTGCGATTTGATC 3541
 QY 2287 AGCCCGCGTGGCGCGCGCATTTAGTGCATGAGTTGGTTCGCGCGCATCAGCTTTA 2346
 Db 3542 AACCTGGCTGGCATCACGAATTTTATCATGGGATCGGCAAGCTGCGTCTCGTTTA 3601
 QY 2347 GCGGTGTTATCACTGAGAGCAGCATCTGCAACAGCTGCTGACCGACGATCTAGCGC 2406
 Db 3602 CTACAGCATATGCGCTGGCGCGGAGTTGATCGTCTGCTGCTGCTGCTGCTGCTGCTG 3661
 QY 2407 TACGATGTCAGCATTTAGGCGGAGCTGAGGCGCGAGCTGAGGCGCGGCTGCGAGCTGCGCG 2466
 Db 3662 AGCGTATGACAAAATTCAGGCGCATTTGCTCTGGCAGGCGGCAACACAGCGCGCGCG 3721
 QY 2467 ACATCTGCGAGCGGCTGTCAGCAGCAGTCTGCTGCGGAGGCGGACCTGATGCG 2526
 Db 3722 ATATGTTGAAACAGGCGATGCG--GACCTGTACGCACTACTCAGTGGGACGATTTATGCA 3780
 QY 2527 CAGCAATACGATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2586
 Db 3781 ACCGCACTATGATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3840
 QY 2587 TCTGCGTCAATTCAGCCACACTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2646
 Db 3841 GCTTCAGCAACAGCATCCGGATATGCGGATCTTGTCTTATTTGAGCGGCTCTGAGCGGG 3900
 QY 2647 AGCAATACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2706
 Db 3901 AGGACCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
 QY 2707 GCAACCGCTGATTTACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2766
 Db 3961 AGCGCCGCTTGTGCTCATCTGCGCGGCTACCGAGTTCGTTTCCCGCAACCGCGTGC 4020
 QY 2767 CAATCTGAGCGGATTTGTTCCATCGCATCAGCGATTTTTCGCGCGCTGCTTTACGC 2826
 Db 4021 CCATGTGAACAGTGGCTACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
 QY 2827 GCGGATGGGTGACGATCTGTGGAACAAACAGCGCTGACAGGTAACAAACAGCGAGCT 2886
 Db 4081 ACAGTTTGGACAAATTTATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4140
 QY 2887 GACCTGGCGGATGCGCTGAACTGCTGCGCAAGTGTGATGATGCTGCTGCTGCTGCTGCTGCT 2946
 Db 4141 CCAGTTAGCGATGGCGGATTTATGACGAGTACAGTGTGATGACGAGCGGCTTACAC 4200
 QY 2947 GCGGACGCAATCTGAGCTGGGTTATCAGGTTTCTTGGACAAAGTGGGAGCTGGC 3006
 Db 4201 GCCTGATTTGCACTACGCTAGGATTCAGGCAATTTATCGGTGAGGAGTGGCACTGAG 4260
 QY 3007 GAGCGGACCGGCTGAGCAGCGGATGCTGATGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3066
 Db 4261 CGCGCGCATGTTTATGCTCAGCGATTTATCGGTGAGGAGTGGCACTTTCGCGATTTATG 4320
 QY 3067 TTATCGTTTGTCTACAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3126
 Db 4321 CTACGCTTTGTTTATACCTGCGCTTTCCGCAACCGCACTGCTGATCGAAGACACACA 4380
 QY 3127 TTACGTTAAACAGCGCGGCTGGCGGAGAACACCGCTGCTGAGCACATCGCGGCTATGC 3186
 Db 4381 CTACATTGCAAGGCTAATCTTTCAGGCGGAGCGGCGGCTGAGAACATTTCCGATTTATG 4440
 QY 3187 CAATCAGCAAGCTGAGCTGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3246
 Db 4441 TCGCGACAGGTTGGCGTTTACAGCCTTGTGCGGAGAACAGGAGTGGATTTGCCAT 4500
 QY 3247 TACCTCGAGCGCAACATCGATTCGATTTGGCAACAGCAGCGCGGCTGAGCGTGCAGCG 3306
 Db 4501 TAGCTTAAAGCGGCTAATCGTCACTTTTGGCAACAGCAAC---CGCAAGCTGTAGCGG 4557
 QY 3307 CCGTGGCGCGGCTGTTTCAATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3366

Db 4558 ATTACGGCGGGCTGTTTTCATCCGACAAACCGGCTACTCCCTACCGCTCCGGTGGCGCT 4617
Qy 3367 AGCGAGTGTGTAGCAGCGCTGTGTGCCACCGATGCCCTCAGCTCAGCAACATATCGA 3426
Db 4618 GGCCGATCGCTCAGCGCGCTGGATGTGTTACCTCTTCTCTGTTTACACAGACGATTGC 4677
Qy 3427 AGCTTTGGCCGTCAGCAGTGGCGGAAACAGCGATTTTTCCGCTGCTGCTAAACCGCATGCT 3486
Db 4678 TCACCTTTGCCAGCAACGTTGGCAGCAACAGGGGTTTTTCCGCATGCTGAATGCGATGTT 4737
Qy 3487 GTTTTTGGCCGTAAGCCGAGCAGCGCTGGCGCGTGATGCAACAGTTTTTACCGGCTCGA 3546
Db 4738 GTTTTTAGCCGACCGCCGAGTCAAGCTGGCGTGATGCGAGCTTCTATGGCTTACC 4797
Qy 3547 TGC CGGGTTAAATPAGCGCTTTTACCGCGGGCAACTGCGCTGCGGATATAAACCGCGAT 3606
Db 4798 CGAGGATTTGATGCGCGCTTTTATCGCGGAAACTCACCGTGACCGATCGGCTACGCAT 4857
Qy 3607 TCTGTGGCGCAAGCCCGCTGCGGCTCCCATCGGTGAGCGCTGCGGCGCTGTTGAATCTGT 3666
Db 4858 TCTGAGCGCAAGCCCGCTTCCCGTTTTTCGCGGCAITTCAGGGCAATATGACGACTCA 4917
Qy 3667 CGAACCAAGGAGAAATAATGAAACGCACTTATGTGATTTGGCGCAGGCTTTGGCGGCTG 3726
Db 4918 TCGTTGAAGCGACTACATGAAACCAACTACGTTAAITTTGGTGGCGCTTTGGTGGCTG 4977
Qy 3727 GCGCTGGCAATGCGCTGCAAGCGGGGATACCAACCACTTACTCGAGCAGCGGAC 3786
Db 4978 GCACTGGCAATTCGTTTACAGGCGCGAGGTATTCCTGTTTGTGCTTTCGAGCAGCGGAC 5037
Qy 3787 AAACCGGGCGAGCGGCTATGTGTTTGAGGACAGTGGCTTTTACCTTCGATGCGGACCC 3846
Db 5038 AAACCGGGTGGCGGCTTATGTTTATCAGGACAGGGCTTTTACTTTTGTATGAGCGGCT 5097
Qy 3847 ACGGTGATCACCGATCCAGCGCATCGAAGAGTTCGTTCACTGCTGGCAGGAAATCGCTC 3906
Db 5098 ACGGTTATCACCGATCCAGCGGATTTGAAGAACTGTTTGTCTTGGCCGCTAAACAGCTT 5157
Qy 3907 AGCGATTACGTAGAGCTGATCGCGTAAACCGCTTCTATCGCTGTGCTGGGAAGATGGC 3966
Db 5158 AAGGATTACGTAGAGCTGTTGCGGTCAGCGGTTTTTATCGGCTGTGCTGGAGTCCGCG 5217
Qy 3967 AAACAGCTTGATTACACAAATATACCGCGCTGCTGGAGCAGCATCCGCACTTCAAT 4026
Db 5218 AAGGTCTTCAATTTACGATTAACGACAGGCGGCTAGTAAGCGCAGATACAGCAGTTTAA 5277
Qy 4027 CGCAAGATGTAGAAGCTATCGTCAATTTCTTGCCCTATTCAGCTGAAGTATTTAGAGAG 4086
Db 5278 CCGCGCATGTTGCGGGTTATCGAGCGTTCCTTGACTATTCGGTGGCGGTTTCAATGAG 5337
Qy 4087 GGTATCTGAAACTCGGCAAGTGGCGGTTTCTGAGGTGCGTGACATGCTGCGCGTCGG 4146
Db 5338 GGTATCTGAAAGCTCGGCACTGTGCGCTTTTATCGTTCAAGACATGCTTCGGGCGCG 5397
Qy 4147 CCGCAGTTGGAGCTCTGCAAGCATGGCGGCTCAGCATGTTGTCGGAATTTATT 4206
Db 5398 CCCCAGTTGGCAAGCTGAGGCAATGGCGAGCGTTTACAGTAAAGTTTCCGCGCTACATT 5457
Qy 4207 CAGGACGATCATCTGCGTCAGGCGTTTTCTTCTCACTCATTTGCTGGTGGCGGTAATCCT 4266
Db 5458 GAGGATGAGCATCTTCGCGAGGCGTTTTCTTTTCACTCGCTTCTAGTGGGGGATCCG 5517
Qy 4267 TTTGCAACGTCATCGATCTATACCTTAAATTCATGCGCTGGAGCGTGAATGGGCGTGG 4326
Db 5518 TTTGCAACCTCGTCCATTTATACGCTGATTCACGCGTTAGAACGGGAATGGGCGTCTG 5577
Qy 4327 TTTCCGCGGGCGGACCGCGGCTGTGGTGCAGGGCATGCGGGACTGTTTCAGGACTTG 4386
Db 5578 TTTTCAAGCGGGTGGAAACCGGTGCGCTGGTCAATGGCATGATCAAGCTGTTTCAGGATCTG 5637
Qy 4387 GCGCGGAGCTGTTACTGAAATGCGAAGTGAGCGGCTGGAACCAAGCGGCAATCGCAT 4446
Db 5638 GCGCGGAAAGTCTGCTTTAACCGCCCGGCTCAGTCATATGAAACCGTTGCGGACAAAGATT 5697

Qy 4447 AGCGCGCTTCACTTAGAGGGCGGACGACGCTTTCGATGCGCGCGCTGTGTGCGCTTCAATGCC 4506
Db 5698 CAGGCGCTGCGAGTTGGAAAGCGGACGAGCGGTTTGAACCTGCGCGGTGCGTTCGAACGCT 5757
Qy 4507 GAGCTGGTGCATACCTACGACAAACGCTTCCGACCACTCCGCTGGCAATGAACGTCGG 4566
Db 5758 GATGTTGTACATACCTATCGCGATCTGCTGTCTCAGCATCCCGAGCGGCTAAGCAGGGG 5817
Qy 4567 ACATCGCTGAAGCGTAAGCGCATGAGCAACTCGCTGTTTGTACTCTATTTTGGCTGAAAT 4626
Db 5818 AAAAAGCTGCAATCCAGCGTATGATTAACCTACCTGTTTGTACTCTATTTTGGTCTCAAC 5877
Qy 4627 CAGCGCGATGAACAGCTCGCGCACACACCGCTGTTTGGCCCGGCTTATCGTAGTTG 4686
Db 5878 CATCATCAGATCAACTCGCCCATCATACCGCTGTTTGGGCGCAGCTACCGTGAACCTG 5937
Qy 4687 ATCGATGAGATTTTCAACAGCAGCAGCGCTGGCAGACGATTTTCTACTTTTACCTGCAACGG 4746
Db 5938 ATTCAGGAAATTTTAAACCATGATGGTCTGGCTGAGGATTTTTCGCTTTATTTTACACGCA 5997
Qy 4747 CCTTCAGCAGCGATCCGCTCGCTGCGACCGCCCGCTGCGGAGCTTTTATGTGTAGCG 4806
Db 5998 CTTTGTGTACCGGATCCGCTACTGGCACCGGAAGGTGCGGAGCTATTTATGTGTGGCG 6057
Qy 4807 CCGGTGCGCATCTCGGCAACCGCTGACATCGATCGATCGCAACAGGAAGGACCGGCTTGGCG 4866
Db 6058 CCGTGTTCACACTTAGGCAACCGGAACTCGATCGGCGGTAGAAAGGACCCGACTGCGC 6117
Qy 4867 GATCGAATTTTGTCTTATCTGAGCAGCACTACATCGCGGATTAAGTCAAGCAATTAGTG 4926
Db 6118 GATCGATTTTGTGACTACCTTGAGCAACATTAATCATGCTGCGTTGCAAGACGAGTTGGTG 6177
Qy 4927 ACACAGAAATTTTACGCGCTTTCATTTTCCGACACGCTGCATGCCCATCACGCGCTCG 4986
Db 6178 AGCAGCGTATGTTTACGCGGTTTCGATTTCCGCGAGAGCTCAATGCGCTGGCAAGTTTCG 6237
Qy 4987 GCGTTTTGCTGGAGCGGATTTTGAACGAAAGCGCTGTTCCGCGCGCATAACCGCGAT 5046
Db 6238 GCGTCTCGGTTGAACCTTATCTGACCCAGAGCGCTGTTCCGACCAATTAACCGCGAT 6297
Qy 5047 GCGGATATCAGCAATCTCTATCTGCTGGTGGTGGTACGATCCAGGCGCGGCGGCTGCC 5106
Db 6298 AAGCAGATGATTAATCTTTATCTGTTGGCGGAGGACCCATCTGCGCGGCGGCTTCCC 6357
Qy 5107 GCGGTGATCGGTTGCGGCAAGCGCACCGCAGGCTGATGCTGGAGGATCGCGCGCAATGA 5166
Db 6358 GCGGTAATCGGCTCGCGNAGCGGACGCGAGGCTTAATGCTGGAGGACCTGATTTGACGA 6417
Qy 5167 ATCGACAGCTTTTACTTGAGCAAGTAAACGAAACCAATGCGGCTGGGCTCGAAGAGTTTCG 5226
Db 6418 AT---ACGTCATTACTGAATCATGCGCTGCAAAACCAATGCGGCTGGCTCGAAGAGCTTTG 6474
Qy 5227 CCACCGCGCCCAAGCTGTTGATGACCGCGCGGAGCAGCTGATGCTGTATGCTT 5286
Db 6475 CACTGCTATCGACGCTTTTCGACGCCAAACCCGTCGCGAGCTGCTGATGCTTACGAT 5534
Qy 5287 GGTGCTGCTACCTCGCATGATGATGATTTGATGCGGCAACCGCTGGGCGAAAGCGGCAACGACG 5346
Db 6535 GGTGCGGCACTGCGACGAGCTCATTTGACGATCAAAACACTGGGCTTTCATGCGGACGAC 6594
Qy 5347 ATGCGCTGGAAGACCGCGGAGCACTGATGACGATCTGCAAAATGAAACCCCGCGGCTT 5406
Db 6595 CCTCTTCGAGATGCTGAGCAGCGCTTGCAGCAGCTTGAATGAACCGCTCAGGCGCT 6654
Qy 5407 ACAGCGGCGGACATGATGAACCGGCTTTAGGCGGTTTACGAGTTCAGGAGTGGCATTTTC 5466
Db 6655 AGCGGCTTTCGAAATGACGAGCGGCTTTTGGCGGCTTTCAGGAGTGGCGG 6714
Qy 5467 ACCAGCTCGCGCAACAACTGCGGCTTGTATCTGGAAGGCTTTCGCTATGATGACGACGA 5526
Db 6715 ATGATATCGCTCCCGCTACGCGTTTCGACCATCTGGAAGGTTTTCGCTGATGATGTCGCG 6774

QY 969 CTTCTTCGCCAGGATTTGGCCAGCGGTTTCAA CTGCTCGAGACCTCGCCGACGGTTG 1028
DB 3044 CTGCTTTGGCCAGGATTTAGGCGAGCGGTTTCAGCTGCTGGACGATCTGGCGGACGGCCA 3103
QY 1029 CAAACACACCGGTAAAGATGTGCACAGGATCAGGCAAAATCCACGCTGCTGATCAGATGCT 1088
DB 3104 TGCCGGGACCGGCAAGACATCAATAAGACCGGGTAAGTCCACGCTGCTGGCGATGCT 3163
QY 1089 CGGTGCTGACGGCGCGGAAACGTGCTGCGCGATCACTGCGAGCGAGATGCACACCT 1148
DB 3164 CGGACGAGCGGGTGGCGAGCGGCTCGACACCCATCTGCGCGCGAGACGCCCATTT 3223
QY 1149 TGCCCTGGCTGCCATCGCGGCATCGCCACTCGCCAAATATATGACGCGCTGTTTAATCA 1208
DB 3224 TTACCGCGCTCGGAAAAAACAGGCCACGCGACGCTTTATGACGCGCTGTTTCAA 3283
QY 1209 ACAGCTAGCGATATCAA CTGAGCGCGGCTCAGCGGTGGGCACCTTTGCGGTGATCGCG 1268
DB 3284 ACAGCTGGCGGTTTGTGCTGAGCAACGGATACACCCCGTAAATTTGTGAGATCACA 3343
QY 1269 CCGCGCTCTACAGCCACTTTTACGCGTTTCAGGGGTTAGCACAAACGCTGCTGGCGCGC 1328
DB 3344 T-----GA 3346
QY 1329 GGCCATCGCATCACATTCATCAGCAAGCGGATGCCCGCACTTTGCTTAGCGACGACGC 1388
DB 3347 AGGACGCGCATCTGGTTGACGCTAAATAATGAACACCTGATATCTGCTGCAACCTTGACC 3406
QY 1389 ATCGATTTTGTGCGCTGCGCAACAGACGATCTCTGCGGTTGCTGCGCGCGGTTG 1448
DB 3407 GGGCGATGATACCATTCGACCGGATTTGA CGCTGGGTTTTGAACTGCG----- 3459
QY 1449 CATCGCTGCGCTCGCGCGGCGCTGCTGCTGTTTTCGGGTGATCGACGATCTCGCGTCC 1508
DB 3460 -----GCCCTCCGAGCTGGATCTCGACGGTATCGATCTCTCCACC 3501
QY 1509 TGCACCGATATGCTGTGCGCGAATCGCTGCGGTACTGAAGCATTTGAACATCGATGCG 1568
DB 3502 ACCCTGTTTTTCCCGCGCGCTGAAGCCCGCGTGTGATCAGCTCCATGACCGCGCGCGG 3561
QY 1569 GTGATCGCGACGAAATGAAGCGGCGGCGGATTTGCTGTAAGCGCTGCTATCGCG 1628
DB 3562 GCGCGCGCCAGACATTAACCGTCACTGCGCCAGCGCGCGCAACCTTTGGGCTGGCG 3621
QY 1629 TTTGTTTCGGTGGCTCGCGCTTTCGCGTCAATCGTGAAGCGGGAATTCGCTTGGGTTG 1688
DB 3622 ATGGGCGTGGTTCCAGCGCGTGGCGCTGGAGACGGCGGCGAGCACGGGCTGGATGCC 3681
QY 1689 ATGCCCTTTCGTTTTCACAGGATGACAAAGCGCTGAACGTTTTCAGGCCAGCAGCAT 1748
DB 3682 CAGCTACGCCATATCGCCCGGACGTGCGGCTGCTG-----GCT 3720
QY 1749 ATCTATGATCGCATATGCTGCTACGCGCGAGCTGATCTCTCAAACACGCGGGGTTT 1808
DB 3721 AACCTTGGCGCGCGAGATCCCGGTGCGAGGGCTGGAATAACCGCGCGCGCGGTG 3780
QY 1809 AATTTGACGAGCGCGCGGATTAATCATCATGCTGCTGCGCGCTGCGGACAAATCAGCCAG 1868
DB 3781 GACATGATCGACGCGAGCGTTAATTTGTGATCTGAACCGGCTG-----CAG 3828
QY 1869 ATGGTGGCGGCTTTGATTTTCAAGTCAAGCAACTGCCCGCTGCTATCAGCGCGTGGGG 1928
DB 3829 GAGGCGCTCCAGGCGCGCGCATCGCGATGCGCGGATCTCTCAACGCCATTTGCGCAG 3888
QY 1929 CCACCTCGCGCGCGGTTTCTCTGCGCGCTCCATGCGCGCTGCGCGAGCGCTGCGTCAG 1988
DB 3889 CTGGTGGCGAATCGCGGTACCGGTGCGGTGTTAAAGAG--GTGGGCGCGGATCTCCC 3946
QY 1989 CCGGTGTTTATGCTTTCGCTGGGTAGCTGCAAGGCCATCGCTTTCGGCTGTTTTCGAT 2048
DB 3947 CGGACGTTGCTGCGCGGAGCGTGGCGGTGGGATGATCGACATTTG----- 3997
QY 2049 CTGGCGAGCGGTGCGCGAGCTGCGGCTATCGCTGGTATCGCCCATTTGTGGGGATTA 2108

DB 3998 CCGCGCGCGCGGAAACACAGCTGGGCGCGGTGGAAGCTGAACCGCGCCCCG----- 4047
QY 2109 AACGCCGCAACACAGCATCAGCTGAGAGCTCGCTGCGCGCGGTGGGTGACGGAATTCGTGTC 2168
DB 4048 ---ACCCCGAGGCGGAAATGTGGCATGGCTTTGCCGACTGGGCAATCTCTACTGCC 4104
QY 2169 GATCAGCGCGCAGCGCTTACAGCACGCGCAGCTGTTTATACATCACTGCGGGTTAAACAGC 2228
DB 4105 GATGGCTGCGTGGCGTCCATCTTGGCTGCCGATATCCCGCTTATCCGCTCCGCGCGC 4164
QY 2229 GCGCTGGAGACTGGAATGCGGTACCGCATGCTGGCGCTGCCGATTTGTTGATCAG 2288
DB 4165 ATGCGCAACGCGATTTGACGCAAA-----AAGCCATCGCGTGGGTGAGATCTG 4215
QY 2289 CCGCGCTGGCGCGCGCATTTGAGTGGCATGAGCTTTGGTGGCGCGGCATCAAGCTTTAGC 2348
DB 4216 GTGGGCGAGCGCGCGGTGCTGGCGCATGCCAACGCCCTCGCGCAGCGGCAATTTGCC 4275
QY 2349 CGTGTTCATCAACTGGAGCAGCATCTGCAACAGCTGCTGACCGACGATCGTTACGCGCTA 2408
DB 4276 CATTTCCGCACTGATTAACGCACTGC-----GGATCGCTGTTTCTGTACCGCGCAG 4328
QY 2409 CGGATGTCAAGCATTTCAAGCGCAGCTGCGAGCGCGCGGTTGCCAGCTGCGCGCGAC 2468
DB 4329 TGCAAACTGCAAGCGGTTCGACACGCGCTGCTTCGCGTCAACGCGCGCGCATCCCT 4388
QY 2469 ATCGTCTGAGCAGCGCTGTCGCGAGCAAGTGTGCTGGCGGAGCGACCTGATCGCGA 2528
DB 4389 GTGACGCGACTAGCGTCCCTTATACCGGGAGCGGT-----ATGA 4427
QY 2529 CGCAATACGATGTGATTTTGGTGGTGTGGA CTGGCGAATGGCTTGTGATGGCTGCGTCC 2588
DB 4428 AAAAAATGGGATCTGATTTCTGGTGGCGGGGCTGGCCAAACGGGCTTATCGCTGCGCAG 4487
QY 2589 TCGGCTCAATTTGAGCAGCACAACTGAATGCTGCTGCTGGAGAGCGATGCGCATCCGCGCAG 2648
DB 4488 TAAAGAGCGGTATCCGACGCTTGTGTTAATGCTGGAGTGGCGGAGCGGCCCGCGG 4547
QY 2649 GCAATCATACCTGCTGTTTTCATACAGCGATCTCAGCGCCGAAACAACTTCGCTGCTGTC 2708
DB 4548 GAAACCAACCTGCTGCTTTCACCAACAGATATCAGCGCCAGCCAGCAGCGCTGCTG 4607
QY 2709 AACCGCTGATTAACGCTGCTGCTGATTAACAGGTGCTGCTGCTGCTGCTGCGCGCA 2768
DB 4608 CGCGCTGCTGCGCCATCGCTGGGACGGGTACAGCTCCACTTTCCGAACGTTGCTGCGCA 4667
QY 2769 ATCTGACGCGGATTTATTTCCATCGCATCAGGCGATTTTCCCGCCATCTTTACGCGG 2828
DB 4668 CCTCGATGACGGCTACCTGACCATCACTCCACGCTTTTCCCAAGGATGCGCGGGC 4727
QY 2829 CGATGGGTGACCATCTGTGGACAAACACAGCGGTACACAGGTAAACCCACGCGAGTGA 2888
DB 4728 TGATGAAGAGATTTTGTGACAAACGTCACCGTGTACAGGGTGAGCGGCGAGGAATA 4787
QY 2889 CGCTGGCGGATGCGCTGAACTTTGTGCGCAAGTGGTGAATGATGCTGCGCGCTGCGCAG 2948
DB 4788 CCTCAGCGACGACGACGCTTTACCGCGGCGGTGATGATGCGCGCGCTATCAGC 4847
QY 2949 CGACGCCCATCTGACGCTGGGTTATCAGGTGTTTCTGGACAAAGATGCGAGCTGGCGC 3008
DB 4848 CCTCGCGCACCTTCAGCATTTGGCTATCAGGCGTTTATCGCGCAGGATGGCACTGACCG 4907
QY 3009 AGCCGCAAGCGCTGACGACGCGATCTGATGGATGCCACCTGATCAGCAAGCGGGTT 3068
DB 4908 CGCCCCACGGGTTAACCGCGCCGATCTGATGATGCTCCGCGCTCGCCAGGGCAACGGCT 4967
QY 3069 ATCGTTTGTCTACAGCTGCGCTCAGCGCGCATGGCTATTGATTGAAGATACCATTT 3128
DB 4968 ACCGCTTTGTCTATACCTTCCGCTCAGCGCGCACACCTGCTTATCGAAGACACGCACT 5027
QY 3129 ACGTTAAACGAGCGCGCTGGCGGAGAACACCGCTGTCAGCACATCGCGCACTATGCCA 3188

QY	2019	CAAGGCCATCGCTTCGGCTGTTTCTGCACTGGCGCAGCGTGC	2078	Db	7367	CCGACACGCTGCTGATCGAGGATACGCGCTACGCCAATGCTCCCGCAGCGTGATGATAATG	7426
Db	6311	CAGGGCCACCGCTACTACGCTGTTTTCAGAAATCGCCCGCGCTGTG	6370	QY	3159	CCGCTGCTGAGCAGATCGCGGCTATGTCNAATCAGCAAGCTGGAGCGCTGAGTACGCTGC	3218
QY	2079	TGCTGTGATCGCCCATTTGGGGATTAACCGCGGAAACAGACGATCAGCTGAGGCTC	2138	Db	7427	CCCTACCCAGACGGTTTACCGACTATGCTCAGCAAGAGGTTGGCAGCTGGCCACGCTTG	7486
Db	6371	GAGGTGACCATTTGCCACATGCGATGGCTGAGCGCCGCCAGCGGCGACTGCTCTACGCC	6430	QY	3219	TGCGTGAAGAGCAGCGGCATATTACCGATTACCTCGAGCGGCAACATCGATTCGATTCGGC	3278
QY	2139	GCTGGCGCGGTGGGTGACGATTTGCTGATCAGCGCGCAGCCCTTACAGCAGCGGAG	2198	Db	7487	AACGAGAGGAGACGGGCTGTCTGCCGATTACCTTGGCGGCTGACATCCAGGCTCTGTGGG	7546
Db	6431	TGCGGCGGACGAGGTGCTGAGCTTTGTCGACCAAGCCCGCTACGTTGCCGAGGCTAAT	6490	QY	3279	AACAGCAGCGCGGCAAGCGCTGAGCGGCTTGGCGCGCGGCTGTTTCATGCGCACCCAG	3338
QY	2199	CTGTTTATCACTCATCGCGGTTAACAGCGGCTGGAAGCACTGGAATGCGGTACGCG	2258	Db	7547	CCGATGCGCGGCGCTGCCGCTCGGGAATGCGGCTGGGCTATTTCACCTTACCCACTG	7606
Db	6491	CTGGTATCACCCACCGCGGTCTCAATACCGTACTGGATGCGTGGCTGCCCGCAGCGCG	6550	QY	3339	GTTACTCTTGGCTCGCTCGCGCTAGCGGCTAGCGAGTTGGTAGCAGGCGCTTTCGCCACCG	3398
QY	2259	ATCTCGCGCTGCCGATGCTTTTGATCAGCCCGCGTGGCGCGCGGCGNATGAGTGGAT	2318	Db	7607	GCTATTTCGCTCGCTCGCGCTGGCGCTTGGCGAGCGGATTTGCCGACAGCCCGGCGTGG	7666
Db	6551	GTGCTGGCGGTGCCACTCTCTTTTCGACCAAGCCCGCGTGGCTGCCGCTGCTATAAC	6610	QY	3399	ATGCCCTCAGCTCAGCAACATATCGAAGCTTTGCGGCTCAGCAGTGGCGGCGAAGCAGC	3458
QY	2319	GAGCTTGGTCGCGCGCATCAGCTTTAGCGGTTCATCACTGGAGCAGCATCTGCAA	2378	Db	7667	GCAGGTTCCGCTCTATCAGCTCACCCGCGAGTTTTCGCAACCGCACTTGGCGCAGGCGAG	7726
Db	6611	GGGCTGGGTGCGCGGTATGCGCTTTGCCAGACAGCAGCAGCTGCGCGATGAGATTGCC	6670	QY	3459	GATTTTTCGCTGCTCTAAACCGCATGCTGTTTTCGCGCGGTAAAGCGCAGCAGCGCTGGC	3518
QY	2379	CAGCTGTACCGACAGCATGTTACGCGCTACGATGTACGCGATTTCAGGCGCAGCTGCAG	2438	Db	7727	GATTTCTTCGCGCTCTGAAACCGGATGCTTTTCTTGGCGCGGCGGAGAGAAACCGCTGGC	7786
Db	6671	CAACTGCTGGGGATGAGAGCTGCATCAGGCTTGGCGAGCGGCCCGCCAGCAGCTTAAC	6730	QY	3519	CGCTGATGCAACGTTTTCACCGGCTCGATGCGCGGTTAATTAGCCGCTTTTTCAGCCCGGC	3578
QY	2439	CGCGCAGCGGTGTGCCAGCGTGGCGCGCATCGTTCGAGCAGCGCTGTGCCAGCGCAA	2498	Db	7787	GGGTGATGACAGCGCTTTATGCGGCTCGCGGAGCCACCGTAGAGCGCTTTTTCAGCCCGTC	7846
Db	6731	GAGCGCGGGGACGCGCGCTGGCGGCGACCTGATTGAACAGGCGCAT-----	6777	QY	3579	AACCTGGCGCTCGCGGATTAACCGCGGATTTCTGTGGCGCAAGCGCGCTGCCATCCGTTG	3638
QY	2499	GTGCTGTGCGGAGGCGACCTGATGCGCAGCAATACGATGATGATTTTGGTTCGGTGTG	2558	Db	7847	GGCTCTCTCTTTTGTGTAAGGCGCGCATTTTTCAGGCGCAAGCCACCGGTTCCGCTGGCGG	7906
Db	6778	-----AGCAGGAGTGAAGCGGTATCGTAGGAGTCTGATTTAGTTCGGCGCG	6826	QY	3639	AAGCGCTCGCGCTGTGTTGA---ATTCTGTGCAACCAAGGAGGAGAAAAATGAAACGCAC	3695
QY	2559	GACTGGCGAATGCTTGAITTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2618	Db	7907	AAGCTGGCGGCGCGCTGAAACCATTTCTCTGACAGCAGATATAAGGATGAAAAAAC	7966
Db	6827	GCCTGGCCAAAGGCTGATCGCTGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG	6886	QY	3696	TTATGTGATTTGGCGCAGGCTTTTGGCGGCTTGGCGCTGCGCTGCGCTGCGCTGCGCTG	3755
QY	2619	TGTTGCTGAGAGGATGCGCATCCGCGAGCATCATACCTGCTGCTGCTGCTGCTGCTG	2678	Db	7967	CGTTGTGATTTGGCGCAGGCTTTTGGTGGCTTGGCGCTGCGCTGCGCTGCGCTGCGCTG	8026
Db	6887	TGCTGATCGAGGCGGGGAGGAGCGCGCGCGGAGCAACATCTGTTATTTCCATGAGAGC	6946	QY	3756	CATACCAACACCTTACTTCGAGCAGCGCGCAAAACCGCGGCGAGCGCGCTATGTGTTGA	3815
QY	2679	ATCTCAGCGCGCAACAACTTCGCTGGCTGCAACCGCTGATTACCGTTCGCTGCTGCTG	2738	Db	8027	GATCCCAACCGTACTGCTGAGCAGCGCGCAAGCGCGCGCTGCGGCTACGCTGCGCA	8086
Db	6947	ATCTGATCTCCGCGCAGCAGCGCTGCTGCGCGCTGCTGCGCGCTGCGCGCTGCGCG	7006	QY	3816	GGACAGTGGCTTTTACTTTTCGATGCGCGACCAACCGGTGATCCCGATCCAGCGCCATCGA	3875
QY	2739	ATCAGGTGCGTTTTCCTGCGCTGCGCGCAATCTGACGCGGATTTATTTGCTCCATCG	2798	Db	8087	TGACCCAGGCTTTTACTTTTACCGCGCGCGCTGATCACCAGTCTCTCCGCGCTTGA	8146
Db	7007	ATGAGTGCAGTTTCCCGATCTTCGCGCTGCGCTGCGCGCGCGCTACTCTCCATACCT	7066	QY	3876	AGAGTTGTTACGCTGGCAGGAAATCGCTCAGCGATTACGTCGAGCTGATGCGGTAAC	3935
QY	2799	CAGCGATTTTGGCGGCTCTTTTACCGCGGATGGGTGACGATCTGTGACAAACACAG	2858	Db	8147	GGCGCTGTTTACCTTGGCGCGCAGCGCATGGAGGATTACGTCAGGCTGCTGCGGTAAC	8206
Db	7067	CAGAGCGTTTTCGAGGCGCTGATCAGGCGCTGGGGGAGACATCTGCTAACTGTT	7126	QY	3936	GCCTTCTATCGCTGCTGGGAAGATGGCAACAGCTTGTATTAACGACAATAATACGCC	3995
QY	2859	CCGTACCAAGGTAAACCCACGAGGTGACGCTGGCGGATGGCGGTGAATCTGCTGCGC	2918	Db	8207	ACCTTCTACCGACTCTGCTGGAGTCCGGGAAGACCTTCGACTATGCTAACACAGCGC	8266
Db	7127	CGGTGAGCGAGGTGTATCCCAATAGCGTGCCTTTCGCAACCGGTGAGGCGCTGCTGCG	7186	QY	3996	GCTGTGGAGCAGCAGATCGCCAGCTTCAATCCGCAAGATGTAGAGGCTATCGTCAATT	4055
QY	2919	AAGTGTGATTTGATGGTTCGCGCTGCGCGCGCAGCGCAGCATCTGCGCTGGTTATCAGG	2978	Db	8267	CGAGCTTGAAGGCGCAGATTACCCAGTTCAACCCCGCGCAGCTGAGGCGCTACCGCGCTT	8326
Db	7187	GAGCGGTGATTTGAGCGAGCGCGCGGTGACCGCGATTTCGCGCATGCAACCGGCTATCAGC	7246	QY	4056	TCCTGCTTATTCAGCTGAAGTATTTAGAGAGGCTTATCTGAAACTTCGGCAGCGCTCCGCTT	4115
QY	2979	TGTTTCTTGACAGAGTGGCAGCTGGCGCAGCGCGCAGCGCTTCGACAGCGCGATCTGA	3038	Db	8327	TCCTGGTTACTCCAGGCGGTATTTCCAGGAGGGAATTTTGGCGCTCGCGCAGCGTCCGCTT	8386
Db	7247	TCTTCTTGGTCAGCAGTGGCGGCTGACACAGCGCCACCGGCTGACCGTACCGATCTGTA	7306	QY	4116	TCCTGAGGTGCTGAGCATGCTCGCGCTCGCGCGCAGTTGGAGCGCTTCGCAAGTATGGC	4175
QY	3039	TGGATGCCACCGTGCATCAGCAAGCGGTTATGTTTGTCTACCGTTCGCGCTCAGCG	3098	Db	8387	CCCTCTCTTTCGCGCATGCTCGCGCGCGCGCGCTGCTTAAAGCTCCAGCGCTGGCA	8446
Db	7307	TGGATGCCACCGTGCAGCAGCGAGGCTATGCTTTGTCTACCGCTGCGCGCTCTCCG	7366	QY	4176	CAGCGCTACAGCATGCTGCGGAAATTTTATTCAGGACGATCATCTGCGCTCAGCGCTTTC	4235
QY	3099	CCGATCGGCTATTGATTGAAGATACCATTAAGTTAACCGCGCGCTGCGCGGAGAAC	3158				

8447 GAGCGTCTACCAAGTCGGTTTCGGCTTTATTGAGGATGAGCATCTGCGGACGGCCTTCTC 8506
4236 CTTTCCATCTATTCTGCTGGTGGCGGTAATCTTTTTCACAGCTCATCGATCTATACCTTAAT 4295
8507 GTTTCACATCCCTGCTGGTAGGCGGCAACCCCTTACCACTCGTCCATCTACACCTGAT 8566
4296 TCATGGCTGGAGCGTGAATGGGCGTGTTTTCGCGCGGGGCGCACCGGCGCGTGT 4355
8567 CCACGCGCTTGAAGCGGAGTGGGGGTCTGGTTCCCTGAGGGCGGACCGGGGCGCTGGT 8626
4356 GCAGGCGATGGCGCGACTCTTCGAGGACTTGGGCGGCGAGCTGTACTGAATGCCGAAGT 4415
8627 GAAAGGATGGTGAAGCTGTTTACCGATCTGGGCGGGAGATCGAACTCAAAGCCGGT 8686
4416 GAGCAGCTGGAAAAACAGCGGCAATCGCAATTAGCGGCGTTCAAGTTAGAGGGCGGACG 4475
8687 CGAAGAGCTGGTGGCGGCGATAACCGCGTAAGCCAGGTCGCGCTGGCGGATGGTCCGAT 8746
4476 CTTCCATGCGCGCTGTGGCTTCCATGCGACGCTGGTGCATACCTACGACAACTGCT 4535
8747 CTTTGACACCGACGCGGTAGCTCGAAACGCTGACGTGGTGAACAACCTATAAAGCTGCT 8806
4536 TCGCCACCACTCCGCTGGCAATGAACCTGCGACATCGCTGAAGCGTAAGCGCATGAGCAA 4595
8807 CGGCGACCAATCCGCTGGGCGAGAGCGGCGGCGAGCGCTGGAGCGCAGAGCATGAGCAA 8866
4596 CTCGCTGTTTGAATCTATTTTGGCTGAATCAGCGCGATGAACAGCTCGCGCACACAC 4655
8867 CTCGCTGTTTGTCTCTACTTCGCGCTGAACACGAGCTCATTTCCAGCTGGGCGACCATAC 8926
4656 CGTCTGTTTGGCGCGGTTATCGTGAGTTGATCGATGAGATTTTCAACAGCAGCGAGT 4715
8927 CATCTGTTTGTGCTCCCGCTACCGGGAGCTGATCGACGAGATCTTTTACCGGAGCGCGCT 8986
4716 GCGACAGATTTTCTACTTTTACTGCAAGCGCTTGCACAGCATCGCTGCGTGCAC 4775
8987 GCGGATGACTTCTCGCTCTACTGCACTCGCCCTGCGATCGATCCCTCGCTCGCGC 9046
4776 GCCCGCTCGGACAGCTTTATGTTAGCGCGGTGCGCATCTCGGACCGCTGACAT 4835
9047 TCCCGCTCGCGCAGCTTCTAGTGTGCGCGCGGTGCGCATCTTTGGCAACGCGCGCT 9106
4836 CCACTGGCAACAGGAAGACCGGCTTGGCGATCGAATTTTGTCTTATCTGAGGACGA 4895
9107 GCACTGGGCGCAGGAGGGGCGGAAGCTGCGCGACCGCATCTTTGACATCTTTGAAGAGCG 9166
4896 CTACATGCGCGATTAAGTACGTCAGCAATTAAGTACACAGCAATGTTTACGCGCTTTGATTT 4955
9167 CTATATGCGCGCTGCTAGCAGCTGTGTGACCCAGCGATCTTTTACCCGCGACATTT 9226
4956 TCGGACACGCTGCATGCCCATCAAGCTCGGCTTTCGCTGGAGCGGATTTTGAAGCA 5015
9227 CCACGACACGCTGGATGCGCATCTGGGATCGGCTTCTCCATCGAGCGCTGCTGACCCA 9286
5016 AAGCGCTGTTCCGCGCGCATAAACCGCATGCGGATATCAGCAATCTCTATCTGTTGG 5075
9287 AAGCGCTGTTCCGCGCGCATAAACCGCATGCGGATATCAGCAATCTCTATCTGTTGG 9346
5076 TCGCGGTACGATCCAGGCGCGGCGTGGCGGCTGATCGGTTCGGGCCAAGGCCACGCG 5135
9347 CGCAGTACTCACCTTGGGCGGGCATTTCTGGCTAGTGGCTCGGCGAAGGCCACCGC 9406
5136 CAGGCTGATGCTGGAGGATCGCGCGCAATGAATCGACAGCCTTTTACTTGAAGCAAGTAAG 5195
9407 CAGCTGATGATTGAGGATC---TGCAATGAGCCAAACCGCGCTGCTTGAACACGCCACG 9463
5196 CAAACCATGCGGTGGGCTCGAAGAGTTTGCACACCGCGCGCAAGCTGTTGATGACCGC 5255
9464 CAGACCATGCGCAACCGGCTCGAAAGTTTGGCACCGCTGCGAAGCTGTTTCGACCGCGC 9523
5256 ACGCGCGGACGACGCTGATGTTGATGCTGTGTGCTGCTCACTCGCATGATGATTGAT 5315
9524 ACCCGCGTAGCGTGTGATGCTCTTACACTGCTGCGCGCACCTGCGATGACGCTCATTTGAC 9583

5316 GGGCAAAACGCTGGCGGAAGGCGGCAAGCGAGCATGCGCTGCAAGACGCGAGGACGCTATG 5375
9584 GACCAGACCCACGCTTTCGCCAGAGGCGCGGCGGAGGAGGAGGCCACCCAGCGCTG 9643
5376 CAGCATCTGCAAAATTGAACACCGCGCTTACAGGGCGGCGACATGATGAACCGCG 5435
9644 GCCCGGCTGCGCACGCTGACCTTGGCGGCTTTGAAGGGGCGGAGATGAGGATTCGGCC 9703
5436 TTTTGGGCGTTTTCAGGAAAGTGGCGATCAATTCACAGCTGCCGCAACCACTGGCGTTTGT 5495
9704 TTCGCTGCTTTTCAGAGGTGGCGTGAACCAAGTATTACGCCCGCATGGCGCTCGAT 9763
5496 CATCTGGAAGGCTTCGCTATGAGTGAACGCAACGAATTAACGAGCTTTCGATGACAG 5555
9764 CACCTCGACGGCTTTGCGATGAGCTGGCTCAGACCGCTATGTCACTTTGAGGATACG 9823
5556 CTGCGTCTACTGCTATCAGTTCGCGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5615
9824 TCGCGCTACTGCTATCAGTTCGCGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 9883
5616 GTGCGCGACGAAGCGGTGCTCGATCAGCGCTCGGATTTAGGACTGGCGTTCCAGCTCACT 5675
9884 GTGCGGATGAGCGGTGCTGATCGCGCTCGATCTGGGCTGGCGTTCCAGCTGACG 9943
5676 AACATTGCGCGCAATTTGTAGAAATGCGGAAATGCTGCTGCTATCTGCGCGCAATCC 5735
9944 AATATCGCGCGGATTAATTGACGATGCGGCTATTGACCGCTGCTATCTGCGCGCGAG 10003
5736 TGGCTCGATCAGCGGCGGATTCGCGCGATACGCTGACCGCAACGATCGTGAGCG 5795
10004 TGGCTGCGAGATGCGCGGCTGACCGCGGAGAACTATGCGCGCGGAGAAATTCGGGCGG 10063
5796 CTCGCGCTCACTGGCAGCGCTTTAGTGGCGAGGCGGAACCTTATCTCACTCGCGCGCA 5855
10064 CTGCGCGGGTGGCGAGCGGCTTATGATGCGCGAGAGCGTACTACATCTCTCTCCAG 10123
5856 TCGGTTTACCGGGTTTACGCTGCGCTGCGGCTGGCGCATCGTACGCGCTCGCGCGTT 5915
10124 GCGCGCTACACGATCTGCGCGCGCTGCGCTGGCGGATCGCCACCGCGCGAGCGTC 10183
5916 TATCGGAAATGCGCTCAAAAGTTTCAGCACCGCGGTGTCAGCGCTGGGATTCAGCGCAG 5975
10184 TACCGGAGATCGGTATTAAGGTAAAGCGCGGAGGAGCGCTTGGGATTCGCCCGCAG 10243
5976 CCGCACAGTAAAGTGAATAAAGTGGCGCTGCTGGTGAAGGGGCGAGTTTGGCGATCACT 6035
10244 CACACGCAAGGTGAATAAATTCGATGCTGATGCGCGCACCGCGGAGGTTATTCGG 10303
6036 TCGCGTGTCTGCTGCTGAACCGCGTTCGCTGCTGTTGGAGGCTCTCTGTTGATTT 6095
10304 GCGAAGACGACGAGGTTGACGCGCGCTTGGCGCGCTTTTGGCAGCGTCCCGTTTAGGCG 10363
6096 TACGTCGCTGAGCTGGCGGCGAGCTGGCTTGCAGCTTATTCAGCGGTGGCGGTAGAGGA 6155
10364 GCGCGCATGAGCTTTCAGCGAGATGCGCTGTAGTTCGGCAGGCTTTCGGGCGGATAA 10423
6156 AACCAAAACGACACGAGCTTTCACCGCGCGCACCGCATGATGATGCGGTGGCGCATGT 6215
10424 AACCGAAGGAGACGAGCTTCCCGCGCGCACCGCGTGGTGCAGCGGTGGCGGACGT 10483
6216 ATAAGCGTTAAGATAGCTTTTGGCGGATATAGCGAAACGCGCAGCGTTGATGACCA 6275
10484 AGAGCGCTTACAGTAGCGCGCGGATCGAGTGAAGGGCGCAGCGCTGATGACCA 10543
6276 GCGCATCTGACCATGAGTAGCGCGCTGCTGCTCATTCGCGCAACCACTCACT 6335
10544 GACCGTGTGTCACCGAGTAGAGGAGGAGTAGAGGAGGAGGAGGAGGAGGAGGAGGAG 10603
6336 GCGCGGCGCATGCTGCTGCAACGACATAAATCAGCAATTCGCGCACTACCGCAACA 6395
10604 GCAGGGGCGCAACGCGCGCTGCGCGCATCAGCGCATAGCCACCCCGGCAACA 10663

Qy	6396	GCACCGCATAAAGATCGTTGAGCTCAAACTTACCGCTGTCGGTTTCATCGTCCGACAGAT	6455
Db	10664	CCACCGCAAGAGATCGTTTGTAGCTCAATACGCCCTTGCCGGGGTATGGTGTGACTCAT	10723
Qy	6456	GCCAGCCCCCATCCCCAACCGTGCATGATGTATTTATGCACACGCGCGCTACGATTTCCA	6515
Db	10724	GCCAGCGCCATCCCCAGCCGTGCTAATGTAGCGTGGGTAAACGCGCGGATGCCCTCCA	10783
Qy	6516	TCACCAACCACGGTTGCCAACAAAGATAA	6542
Db	10784	TGCGCATTAACCGTCAAGATGACGATTA	10810

RESULT 9

```

US-10-474-536-45
; Sequence 45, Application US/10474536
; Publication No. US20040176570A1
; GENERAL INFORMATION:
; APPLICANT: Bacher, Adelbert
; APPLICANT: Rohdich, Felix
; APPLICANT: Adam, Petra
; APPLICANT: Amslinger, Sabine
; APPLICANT: Eisenreich, Wolfgang
; APPLICANT: Hecht, Stefan
; TITLE OF INVENTION: INTERMEDIATES AND ENZYMES OF THE NON-MEVALONATE ISOPRENOID
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: 9286.24
; CURRENT APPLICATION NUMBER: US/10/474,536
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: PCT/EP02/04005
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: DE 102 01 458.2
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: DE 101 55 084.7
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: DE 101 30 236.3
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: DE 101 18 166.3
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
; LENGTH: 7494
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: pBScarola
US-10-474-536-45

```

2790	QY	CAATCGCATCAGCGGATTTTTCGCCGCCATCTTTACCGCGCGATGGGTGACGATCTCTGGG	2849
3513	Db	GTATTAATCTCTCAGCGGTTTCGTGTAGGTTTATACAGCGACAGTTTGGCCGCACTTCTGGGA	3572
2850	QY	CAAAACAGCCGCTACAAACAGGTAAAAACCAACGACAGTGCAGTGGCGGATGGCCGTGAAC	2909
3573	Db	TGGATACCGGGTCGCAGAGGTTAATGCGGAATCTGTTCCGTTGAAAAGGTCAGGTTA	3632
2910	QY	TTGCTGCGCAAGTGTGTATGTAGTGTGGCGGCTGACCGGACGCCACATCTCTCAGCTGG	2969
3633	Db	TGGTCCCGCGCGGTGTATGTACCGCGCGGGTTATCGGCAAAATTCAGCACTCAGCGTGG	3692
2970	QY	GTTATCAGTGTGTTCTTGTCGAAGAAGTGTGCAGCTGCGCAGCGCGCACGGCTTCAGCAGC	3029
3693	Db	GCTTCAGCGGTTTATGCGCAGGAATGGGATTGAGCCACCCGCAATGTTTATCGTCTC	3752
3030	QY	CGATCTGTATGATGCCACCGTCGATCAGCAAGCGGGTTATCGTTTTGTCTACACGCTGC	3089
3753	Db	CCATTATCATGGATGCCACGGTCGATCAGCAAAATGGTTATCGCTTCGTGTACAGCCTGC	3812
3090	QY	CGCTCAGCGCGATCGGCTATTGATTGAAGATACCAATTACGTTAACCAAGCCCGCGCTGG	3149
3813	Db	CGCTCTCGCGCACCAAGTTGTTAAATTGAAGACACGCACTATATTGATAATTGCGACANTAG	3872
3150	QY	CGGAGAAACCGCTCGCTCAGCACATCCGCACTATGCCAAATCAGCAAGGCTGGACGCTGA	3209
3873	Db	ATCCTGAATCGCGCGGCAAAATATTGCGACTATGCGCGCAACAGGTTGSCAGCTTC	3932
3210	QY	GTACGCTGCTGTGAAGAGCAGCGGATATTACCGATTACCTGAGCGGGCAACATCGATC	3269
3933	Db	AGCACTGCTGCGAGAGAACAGGGCGCCTTACCCATTACTCTGTGCGGCAATGCCGACG	3992
3270	QY	GATTTGCGCAAACAGCAGCGCGGCAAGCGTGCAGCGGCTGCGCGCGGGCTGTTTTCATG	3329
3993	Db	CAITCTGG---CAGCAGCGCCCTGCGCTGTAGTGGATTACGTGCGGTCGTCTCCATC	4049
3330	QY	CCACCAACGGTTACTCCTTGCCTGCGCGTGCGCTAGCGGAGTTGGTAGCAGCGCTGT	3389
4050	Db	CTACCAACCGGCTATTCACTGCGCTGCGCGGTTGCCCGTGGCCGACCGCTGAGTGCATCTG	4109
3390	QY	TGCCCAACGATGCCCTCAGCTCAGCCAAACATATCGAAGCTTTGCCCCGTGACAGTGGC	3449
4110	Db	ATGCTCTTACGTCGGGCTCAATTCACTAGCCATTAGCAATTTTGCCCGGAGCGCTGGC	4169
3450	QY	GCAAACAGCGATTTTTCCGTCTGCTAAACCGCATGCTGTTTTGGCCGGTAAAGCCGACG	3509
4170	Db	AGCAGCAGGGCTTTTTCCGATGCTGAATCGCATGCTGTTTTAGCCGACCCGCGGATT	4229
3510	QY	AGGCTGGCGGTGATGCAACGTTTTTACCGGCTCGATGCCGGTTAAATTAGCCGCTTTT	3569
4230	Db	CAGGCTGGCGGGTTATGACAGGTTTTTATGTGTTTACTGAAGATTAAATTGCCCGGTTTT	4289
3570	QY	ACGCCGGCAACTGCGCCTGCGCGATAAAACGCGGATTCTGTGCGGCAAGCCCGCGTGC	3629
4290	Db	ATGCGGGAAACTCACTCGCTGACCGATCGGCTACGTTCTGAGCGGCAAGCCGCTGTTT	4349
3630	QY	CCATCGGGTGAAGCGCTGGCGCGCTGTTGAATTCTGTGCAACAGGGAAGAAAAATGAA	3689
4350	Db	CGGTATTAGCAGCATTTGCAAGCCATTATGACGACTCATCGTTAAAGAGCGACTACATGAA	4409
3690	QY	ACGCATTTATGATTGGGCGAGGCTTTGCGGCTTGGCGCTGCGGATTCGCTGCAAGC	3749
4410	Db	ACCAACTACGGTAAATTGGTGCAGGCTTTGGTGGCTTGGCACTGGCAATTCGTCTACNAGC	4469
3750	QY	GGCGGGCATACCAACCACTTACTCGAGCAGCGCGCAAAACCGGGCGGACGCGCTATGT	3809
4470	Db	TGCGGGGATCCCGCTCTTACTGTCTTGAACAAACGTGATAAACCGCGGGCTCGGGCTTATGT	4529
3810	QY	GTTTGGAGCAGTGGCTTTTACCTTCGATGCCGACCCACGGTGCATCACCGATCCCAGCGC	3869
4530	Db	CTACAGGATATCAGGGGTTTACCTTTGATGAGCGCCGACCGGTTATCACCGATCCAGTGC	4589
3870	QY	CATCGAAGAGTGTGTTCACTGCTGCGCAGGAAAAATCGCTCAGCGATTACGTGAGCTGATGCC	3929

[illegible]

5670 TGATTTTCGGGACACAGCTTAATGCCCTATCATGGCTCAGCCCTTTTCTGTGGAGCCCGTTCT 5729

5010 GAGCGAAAGCGCTGGTTCCGCCCGCATAAACCGCGATGCCGATATCAGCAATCTCTATCT 5069

5730 TACCAGAGCGCTGGTTTCGGCCGATTAACCGGATAAACCATTTACTTAATCTCTACCT 5789

5070 GGTGGGTGCGGTACCGATCCAGCGCGCGGCGTGCCTCGCGTGATCGGTTTCGGCCAAAGC 5129

5790 GGTTCGGCGCAGGACACGATCCCGCGCAGGCATTCCTGCGCTCATCGGCTCGGCAAAAGC 5849

5130 CACGGCCAGCTGATCTCTGGAGGATCGCCCGNAATGAATCGACGCTTTACTTTGAGCAA 5189

5850 GACAGCAGGTTTGATGCTGGAGATC---TGATTTGAATTAATCCGCTGTTACTCAATCAT 5906

5190 GTAAACCAACCATGCGGTGGGCTCGAAGAGTTTCGCCACCGCCGCCCAAGCTGTTGAT 5249

5907 CGCGTCGAAACGATGCGAGTTGGCTCGAAAGTTTTCGACAGCCTCAAAGTTATTGAT 5966

5250 GACCCGAGCGCGCAGCACGCTGATGCTGTATGCGTGTGTGCTCACTTCGCGATGATGG 5309

5967 GCAAAACCGCGCGCAGCTACTGATGCTCTACGCGTGGTCCGCCATTTGTGACGATGTT 6026

5310 ATTTGATGGCCAAACGCTGGCGAAGCGGSCACGACGATGCTCGTCGAGACCGCAGGCA 5369

6027 ATTTGACGATCAGACGCTGGGCTTTTCAGGCGCGGACGCTGCCCTTACAAACGCCCGCAACAA 6086

5370 CGTATCGACGATCTGCAAAATGAAACCCCGCGCTCAGCGCGCGCACATGATGAA 5429

6087 CGTCTATGCACTTGAGATGAACCGCCAGGCTTATGCAGGATCGCAGATGCAGAA 6146

5430 CCGCGCTTTAGGCGCTTCAGGAAGTGGCGATCATTTACCAGCTGCGGCAACAACTGGCG 5489

6147 CCGCGCTTTGCGGCTTTTCAGGAAGTGGCTATGGCTCATGATATCGCCCCGCTTACGCG 6206

5490 TTTGATCATCTGGAAGGCTTCGCTATGATGACACGACGAAACATTAACGCGAGCTTCGAT 5549

6207 TTTGATCATCTGGAAGGCTTCGCGATGGAATGTAACGGAAGCGCAATACAGCCAACTGGAT 6266

5550 GACACGCTGGTTACTGCTATCACGTCGCGCGCTGGTTCGGTTTGATGATGCGCGCGTA 5609

6267 GATACGCTGGCTATTGCTATCACGTTGACGGGTGTGCGCTTGATGATGGCGCAATC 6326

5610 ATGGGCGTCGCGAAGAAAGCGGTGCTCGATCACGCTCGCATTTTAGACTGGCGTTCCAG 5669

6327 ATGGGCGTCGCGGATAAACCCACGCTGGACCGCGCTGTGACCTTTGGGCTGSCATTTTCAG 6386

5670 CTCACATAATTCGCGCGGACATTTGTAAGATGCGGAAATATGGTCGCTGCTATCTGCCG 5729

6387 TTGACCAATATTGCTCGCGATATTGTGACGATGCGCATGCGGCGCGCTGTTATCTGCCG 6446

5730 CAATCTGCTCGATCAGCGCGGATTAACGCGCGGATACGCTGACTGACCGCAACATCGT 5789

6447 GCAAGCTGGCTGAGCATGAAGTCTGAACAAAGAAATTTATGCGGACCTGAAACCGT 6506

5790 GACAGCGCTCGCTCACTGGCAGCGGCTTTAGTGGGAGCGCGGAACCTTATCATCTCG 5849

6507 CAGCGCTGAGCGGATCGCCCTCGTTTGGTTCAGGAAGCAGAACCTTACTATTGTCT 6566

5850 GCGCGATCGGTTTACCGGGTTTACGCTGCGCTGCGGTGGGCGCATTCGCTACGCTCGC 5909

6567 GCCACAGCGCGCTGCGAGGGTTGCCCTTCGCTTCGCGCTGGGCAATCGCTACGCGGAAG 6626

5910 GCGGTTTATCGGAAATTGGCGTCAAAGTTACAGCGCGGTGTCAGCGCTGGGATTTCA 5969

6627 CAGGTTTACCGGAAATAGGTGTCAAAGTTGAACAGGCGCGGTACGAAAGCTTGGGATCAG 6686

5970 CCGCAGCGCACAGTAAAGGTGAAAACTGGCGCTGCTGGTGAAGGGGCGAGTTTGGCG 6029

6687 CGG-CAGTCAACGACACGCGCCGAAAAATTAACGCTGCTGCTGGCGGCTCTGTCAGGCC 6746

6030 ATCACTTCGCGTGTGCTGCTCTGAAACCGCGCTCGGCTGGTCTGTGGCAGCGTCTC 6087

6747 CTTACTTCCGGATCGGGCTCATCTCCCGCGCTTCGCGCATCTCTGGCAGCGCCGCG 6804


```

; FILE REFERENCE: CL2360 US NA
; CURRENT APPLICATION NUMBER: US/10/808,979
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US 60/471,904
; PRIOR FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Enterobacteriaceae strain DC260
US-10-808-979-7

Query Match          21.2%; Score 1482; DB 18; Length 1482;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1482; Conserved 0; Mismatches 0; Indels 0; Gaps 0;

QY 3685 ATGAACGCACTTATGTGATTGGCGCAGGCTTTGGCGGCTGCGCTGGCGGATTGCGCTG 3744
DB 1 ATGAACGCACTTATGTGATTGGCGCAGGCTTTGGCGGCTGCGCTGGCGGATTGCGCTG 60

QY 3745 CAAGCGCGGGCATACCAACCACTTACTCGAGCAGCGCAAAACCGGCGGAGCGGCC 3804
DB 61 CAAGCGCGGGCATACCAACCACTTACTCGAGCAGCGCAAAACCGGCGGAGCGGCC 120

QY 3805 TATGTGTTGAGGACAGTGGCTTTACCTTCGATGCGGACCCACCGGTGATCCAGATCCC 3864
DB 121 TATGTGTTGAGGACAGTGGCTTTACCTTCGATGCGGACCCACCGGTGATCCAGATCCC 180

QY 3865 AGCGCCATCGAAGAGTTGTTACGCTGCGGAGGAAATCCTCAGCGATTACGTCGAGCTG 3924
DB 181 AGCGCCATCGAAGAGTTGTTACGCTGCGGAGGAAATCCTCAGCGATTACGTCGAGCTG 240

QY 3925 ATGCCGGTAACGCCCTTCTATCGCTGTGCGGAAAGATGGCAACAGCTTGATTCAGAC 3984
DB 241 ATGCCGGTAACGCCCTTCTATCGCTGTGCGGAAAGATGGCAACAGCTTGATTCAGAC 300

QY 3985 AATAATCAGCGCTGCGGAGCAGAGATGCGACGTTCAATCCGCAAGATGTAGAAGGC 4044
DB 301 AATAATCAGCGCTGCGGAGCAGAGATGCGACGTTCAATCCGCAAGATGTAGAAGGC 360

QY 4045 TATCGTCAATTTCTTGCCTATTACGCTGAAGTATTTAGAGAGGTTTATCTGAACTCGGC 4104
DB 361 TATCGTCAATTTCTTGCCTATTACGCTGAAGTATTTAGAGAGGTTTATCTGAACTCGGC 420

QY 4105 ACGGTGCGGTTTCTGACAGTGCGTACATGCTGCGGCTGCGCGCGCAGTTGGAGCTCTG 4164
DB 421 ACGGTGCGGTTTCTGACAGTGCGTACATGCTGCGGCTGCGCGCGCAGTTGGAGCTCTG 480

QY 4165 CAAGCATGGCGAGCGCTTACAGCATGGTGGCGAAATTTATTTCAGGACGATCATCTGCGT 4224
DB 481 CAAGCATGGCGAGCGCTTACAGCATGGTGGCGAAATTTATTTCAGGACGATCATCTGCGT 540

QY 4225 CAGCGGTTTCTTCCATCATTTGCTGGTGGCGGTAATCTTTTGCACAGTCATCGATC 4284
DB 541 CAGCGGTTTCTTCCATCATTTGCTGGTGGCGGTAATCTTTTGCACAGTCATCGATC 600

QY 4285 TATACCTTAATTCATGCGCTGAGCGTGAATGGGCGTGTGGTTTCCGCGCGCGGCACC 4344
DB 601 TATACCTTAATTCATGCGCTGAGCGTGAATGGGCGTGTGGTTTCCGCGCGCGGCACC 660

QY 4345 GGGCGCTGGTGCAGGCGCATGGCGCACTGTTTCGAGGACTTGGGCGGCGAGCTGTTACTG 4404
DB 661 GGGCGCTGGTGCAGGCGCATGGCGCACTGTTTCGAGGACTTGGGCGGCGAGCTGTTACTG 720

QY 4405 AATGCCGAAGTAGCAGCTGGAACACAGCGGCAATCGCAATTAGCGCGTTTCAAGTAGAG 4464
DB 721 AATGCCGAAGTAGCAGCTGGAACACAGCGGCAATCGCAATTAGCGCGTTTCAAGTAGAG 780

QY 4465 GCGGCGAGCGCTTCGATGCGCGCTTCCAAATGCGCGCTGTCGCTTCCAAATGCGCGCTTC 4524
DB 781 GCGGCGAGCGCTTCGATGCGCGCTTCCAAATGCGCGCTGTCGCTTCCAAATGCGCGCTTC 840

4525 GACAACTGCTTCGCCACCATCCGCTGCAATGAAACGTCGACATCGCTGAAGCGTAAG 4584
841 GACAACTGCTTCGCCACCATCCGCTGCAATGAAACGTCGACATCGCTGAAGCGTAAG 900
4585 CGCATGAGCAACTCGCTGCTTTGTACTCTATTTTGGCTGAAATCAGCCGATGAACAGCTC 4644
901 CGCATGAGCAACTCGCTGCTTTGTACTCTATTTTGGCTGAAATCAGCCGATGAACAGCTC 960
4645 GCGCACACACCGCTGCTTTTGGCGCGGTTATCGTGAGTTGATCGATGAGATTTTCAAC 4704
961 GCGCACACACCGCTGCTTTTGGCGCGGTTATCGTGAGTTGATCGATGAGATTTTCAAC 1020
4705 AGCAGCAGCTGGCAGACGATTTTTCATTTTACCTTGCACGCGCCCTGCGAGCAGCATCG 4764
1021 AGCAGCAGCTGGCAGACGATTTTTCATTTTACCTTGCACGCGCCCTGCGAGCAGCATCG 1080
4765 TCCTGCGCACCGCGCTGCGGAGCTTTTATGTTTGTAGCGCGGTGCCGATCTCGGC 4824
1081 TCCTGCGCACCGCGCTGCGGAGCTTTTATGTTTGTAGCGCGGTGCCGATCTCGGC 1140
4825 ACCGCTGACATCGACTGCGACAGGAGCGCGCTTGCAGATTCGAATTTTTCCTTAT 4884
1141 ACCGCTGACATCGACTGCGACAGGAGCGCGCTTGCAGATTCGAATTTTTCCTTAT 1200
4885 CTGAGCAGCACTACATCCCGGATTACGTCAGCAATTTAGTGACACAGCAATTTTACG 4944
1201 CTGAGCAGCACTACATCCCGGATTACGTCAGCAATTTAGTGACACAGCAATTTTACG 1260
4945 CGTTTGTATTTTGGCAACCGCTGATCCCATCAAGCTCGGCTGCGGTTTTCGCTGAGCGG 5004
1261 CGTTTGTATTTTGGCAACCGCTGATCCCATCAAGCTCGGCTGCGGTTTTCGCTGAGCGG 1320
5005 ATTTTGCAGCAAGCGCTGTTCCGCGCGCATACCGCGATGCGGATCAGCAATCTC 5064
1321 ATTTTGCAGCAAGCGCTGTTCCGCGCGCATACCGCGATGCGGATCAGCAATCTC 1380
5065 TATCTGTGGTGGTGGCTGACATCCAGCGCGGCGCTGCCGCGGATCGGTTTCGCGGC 5124
1381 TATCTGTGGTGGTGGCTGACATCCAGCGCGGCGCTGCCGCGGATCGGTTTCGCGGC 1440
5125 AAGGCCACGCGCAGCTGATGCTGAGGATCGCGCGGATGA 5166
1441 AAGGCCACGCGCAGCTGATGCTGAGGATCGCGCGGATGA 1482

RESULT 12
US-10-808-979-3
; Sequence 3, Application US/10808979
; Publication No. US20040268439A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Juan
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2360 US NA
; CURRENT APPLICATION NUMBER: US/10/808,979
; CURRENT FILING DATE: 2004-03-25
; PRIOR APPLICATION NUMBER: US 60/471,904
; PRIOR FILING DATE: 2003-05-20
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Enterobacteriaceae strain DC260
US-10-808-979-3

Query Match          18.3%; Score 1278; DB 18; Length 1278;
Best Local Similarity 100.0%; Pred. No. 0;

```


QY 2942 CTGACGCCGACCCACATCTGACAGCTGGGTTATCAGGTGTTCTTGACAAAGAGTGGCAG 3001
|
|
|
Db 421 CTGACGCCGACCCACATCTGACAGCTGGGTTATCAGGTGTTCTTGACAAAGAGTGGCAG 480
|
|
|
QY 3002 CTGCGCCAGCCGACCGCTGACAGCAGCCGATCTGATGATGATCCGCTGATCAGAA 3061
|
|
|
Db 481 CTGCGCCAGCCGACCGCTGACAGCAGCCGATCTGATGATGATCCGCTGATCAGAA 540
|
|
|
QY 3062 GCGGGTTATCGTTTCTACACGCTGCGCTCAGCGCCGATCGGCTATTGATTGAAGAT 3121
|
|
|
Db 541 GCGGGTTATCGTTTCTACACGCTGCGCTCAGCGCCGATCGGCTATTGATTGAAGAT 600
|
|
|
QY 3122 ACCATTAGCTTAACAGCCCGCTGGCGGAGAACACCGCTGTCAGCACATCGCCGAC 3181
|
|
|
Db 601 ACCATTAGCTTAACAGCCCGCTGGCGGAGAACACCGCTGTCAGCACATCGCCGAC 660
|
|
|
QY 3182 TATGCCAATCAGCAAGCTGGAGCTGAGTACGCTCTGCTGGTGAAGAGCAGCGCATATTA 3241
|
|
|
Db 661 TATGCCAATCAGCAAGCTGGAGCTGAGTACGCTCTGCTGGTGAAGAGCAGCGCATATTA 720
|
|
|
QY 3242 CGGATTACCTCTGAGCGGCAACATCGATCGATTCTGGCAACAGCAGCGCGCCAAAGCTGC 3301
|
|
|
Db 721 CCGATTACCTCTGAGCGGCAACATCGATCGATTCTGGCAACAGCAGCGCGCCAAAGCTGC 780
|
|
|
QY 3302 AGCGGCTCTGGCGCCGGCTGTTTCATGCAACACCGCTTACTCTTGCCTCGCCGCTG 3361
|
|
|
Db 781 AGCGGCTCTGGCGCCGGCTGTTTCATGCAACACCGCTTACTCTTGCCTCGCCGCTG 840
|
|
|
QY 3362 GCGCTAGCGGAGTTGCTAGCAGCGCTGTTGCCACCGATGCGCTCAGCTCAGCCAAAT 3421
|
|
|
Db 841 GCGCTAGCGGAGTTGCTAGCAGCGCTGTTGCCACCGATGCGCTCAGCTCAGCCAAAT 900
|
|
|
QY 3422 ATCGAACGCTTTGCCGCTCAGCAGTGGCGGCAACAGCGATTCTTCCGCTGCTGCTAAACCGC 3481
|
|
|
Db 901 ATCGAACGCTTTGCCGCTCAGCAGTGGCGGCAACAGCGATTCTTCCGCTGCTGCTAAACCGC 960
|
|
|
QY 3482 ATGCTGTTTTGGCGGTTAAGCGGAGCGAGCAGCGCTGGCGGTGATGCAAGCTTTTACCGG 3541
|
|
|
Db 961 ATGCTGTTTTGGCGGTTAAGCGGAGCGAGCAGCGCTGGCGGTGATGCAAGCTTTTACCGG 1020
|
|
|
QY 3542 CTGCGATCGCGGTTAATAGCGCTTTTACGCGCGGCAACTGCGCTGCGGATAAAGC 3601
|
|
|
Db 1021 CTGCGATCGCGGTTAATAGCGCTTTTACGCGCGGCAACTGCGCTGCGGATAAAGC 1080
|
|
|
QY 3602 CGGATTCTGTGCGCAAGCGCGCTGCGGCTGCGGATGAGCGCTGCGCGCTGTTGAAT 3661
|
|
|
Db 1081 CGGATTCTGTGCGCAAGCGCGCTGCGGCTGCGGATGAGCGCTGCGCGCTGTTGAAT 1140
|
|
|
QY 3662 TCTGTGCAACAGGGAAGAAATGA 3688
|
|
|
Db 1141 TCTGTGCAACAGGGAAGAAATGA 1167
|
|
|

RESULT 14

US-10-804-677-7
; Sequence 7, Application US/10804677
; Publication No. US20040224383A1
; GENERAL INFORMATION:
; APPLICANT: E.I. duPont de Nemours and Company, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2346 US NA
; CURRENT APPLICATION NUMBER: US/10/804,677
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: US 60/468,596
; PRIOR FILING DATE: 2003-05-07
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 7
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Pectobacterium cypripedii DC416
US-10-804-677-7

Query Match 15.5%; Score 1085.2; DB 18; Length 1482;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1234; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

QY 3685 ATGAACGCACTATGTGATGCGCAGAGCTTTGGCGGCTGGCGGTGCGGATTCGCCCTG 3744
|
|
|
Db 1 ATGAACGCACTATGTGATGCGCAGAGCTTTGGCGGCTGGCGGTGCGGATTCGCCCTG 60
|
|
|
QY 3745 CAAGCGGCGGCGATACCAACACCTTACTCGAGCAGCGGACAAACCGGCGGACGCGC 3804
|
|
|
Db 61 CAAGCGGCGGCGTGGCGGTACGCTGCTGGAACAGCGGATAAGCTTGGCGGCGCGC 120
|
|
|
QY 3805 TATGTTTGTAGACAGTGGCTTTTACCTTTCGATGCGGACCCACCGTGTATCCCGATCCC 3864
|
|
|
Db 121 TATGTTTATCAGGATCAGGGTTTTTACCTTTGATGCGGTCGCGGTGATTACCGATCCC 180
|
|
|
QY 3865 AGCGCCATCGAAGAGTTGTTACGCTGGCAGGAAAATCGCTCAGCGATTCAGTTCAGCTG 3924
|
|
|
Db 181 AGCGCTATCGAGCGCTGTTTACGCTGGCAGCAAGCAACTCAGTGTATTATGTGACCTG 240
|
|
|
QY 3925 ATGCGGTAACCGCCCTTCTATCGCTGTGCTGGGAAGATGGCAACAGCTTGTATTACGAC 3984
|
|
|
Db 241 ATGCGGTAACCGCAATTTATCGCTGTGCTGGGAAGAGCGGAGCAGCTGACTACGAC 300
|
|
|
QY 3985 AATAATACGCGCTCTGAGCAGCAGATCGCACGTTCAATCCGCAAGATAGAAAGC 4044
|
|
|
Db 301 AACAAATCAGGCGCAGCTGGAGCAGCAGATTCGCACTTTTATCCCGAGATGTCGCGGT 360
|
|
|
QY 4045 TATCGTCAATTTCTTCCTTATTCAGTGAAGTATTAGAGAGGTATTCTGAACTCGCGC 4104
|
|
|
Db 361 TACCGCCAGTTTCTGCGCTATTTCACAGGATGTGTTTCGTGAGGGGTATCTGAAACTGGGC 420
|
|
|
QY 4105 ACGGTGCGGTTTCTCAGGTGCTGACATGCTGCGCGTGGCGGCGAGTTGGGAGCTG 4164
|
|
|
Db 421 ACCGTACCTTTCTGCAATTTCCGCGACATGCTGCTGCGGCGCACAGCTGGGTGGCTG 480
|
|
|
QY 4165 CAAGCATGCGCAGCGCTCTACAGCATGTGTCGCAATTTTATTCAGGACGATCATCTGGT 4224
|
|
|
Db 481 CAGGCTTGGCGCAGTGTCTACAGCATGTGTCGCAATTTTATTCATGACGATCATCTGGC 540
|
|
|
QY 4225 CAGGCTTTTCTTCCACTCATGCTGCTGGCGGTAACTCTTTGCAACGTCATCGATC 4284
|
|
|
Db 541 CAGGCTTTTCTTCCACTCATGCTGCTGGCGGTAACTCTTTTGCACGCTTTCGATC 600
|
|
|
QY 4285 TATACCTTAAATTCATGCTGAGCGTGAATGGGCGGTGTGTTTCCGCGCGGCGCAC 4344
|
|
|
Db 601 TATACCTTAAATTCAGCACTGGAGCGCAATGGGCGGTGTGTTTCCGCGCGGCGTACC 660
|
|
|
QY 4345 GCGCGCTGTGTCAGGCGCATGCGCGACTGTTCGAGGACTTGGGCGGCGAGCTGTACTG 4404
|
|
|
Db 661 GGTGCGCTGGTTGATGGCATGCGCGCTGTTTCGCGATTTTGGGCGGTGAACCTGCTC 720
|
|
|
QY 4405 AATCGGAAGTGAAGCGAGCTGGAACACAGCGCATTCGATTCAGGCGGTTCAGTTAG 4464
|
|
|
Db 721 AACCGGAAGTGAAGCGAGCTGGAACACAGCGGTAAACCGCATCAGCGGTGTCCAGCTGAAG 780
|
|
|
QY 4465 GCGGACGACGCTTCGATGCGCGCTGTGCGCTTCCAAATGCGCGCTGCTGATACCTAC 4524
|
|
|
Db 781 GATGGCGCGCTTTTTCGCGCGCGCGCTTGGGTCAATGCTGAGCTGCTGATACCTAC 840
|
|
|
QY 4525 GACAAACTGTTCGCGCACCATTCGCTGGCAATGAACCGTGGCAATCGCTGAGAGGTGAAG 4584
|
|
|
Db 841 GATCGCTGTTTAAAGCCAGCATCTCTGCGCGCGGTAAACCGCGGCAACGCTGAAAGCGCAAG 900
|
|
|
QY 4585 CGCATGACCAACTCGCTGTTTGTACTCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4644
|
|
|
Db 901 CGCATGACCAACTCGCTGTTTGTACTCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

Qy	5726	GC	CGCAATCTGCTCGATCAGCGGGATTACGCGCGATACGCTGACTGACCGCAACA	5785
Db	3765	GC	TGCGGACTGGCTGGCGAGCGGG-----GGCGACGGTTGAGGGTCCGGTGCC	3715
Qy	5786	TC	GTGACGGCTCGCCTCACTGCGCGCGTTTAGTGGCGAGCGGAACCTATTATCA	5845
Db	3714	TT	CGGACGGCTTATTCCGTCAATCGCCTGCTTGACGGCGCGAGCCCTATTATGC	3655
Qy	5846	CT	CGCGCGATCCGGTTTACCGGGTTTACCGCTGCGCTCGGCTGGGCCATCGCTACGGC	5905
Db	3654	CT	CGCGCGCGAGGGCTTCGGCATCTGCGCGCGCTGCGCTGCTGATCGCGCCGC	3595
Qy	5906	TC	CGCGGCTTATCGGAAATTGGCGTCAAAGTTGACGACCGCGGTGTCACGCTGGGA	5965
Db	3594	GT	GCGTATCTATCGCGCAATCGGACGCGCATCCGGCAGGGTGGCCCGAGGCCTATCG	3535
Qy	5966	TT	CACGGCAGCGCACCAAGTAAAGGTGAACTGGGCGCTGCTGTTGAAAGGGCGAGTTT	6025
Db	3534	CC	AGCGGATCAGACACGTGAGGCTGCCAAGATCGGGCTTCTGGCGCGGAGGCTTGA	3475
Qy	6026	GG	CGATCACTTCGCGTGTCTGTCCTGAAACCGCGTCCGGCTGCTGTGGCAGCGTCC	6085
Db	3474	CG	CGCGCGCATCGCGCTGCGCGCGCGGAAATCAGCCGACGGCTGTGGACCGACC	3415
Qy	6086	TC	G	6088
Db	3414	G	G	3412

Search completed: January 26, 2005, 02:16:28
Job time : 3368 secs

This Page Blank (uspto)

Result No.	Score	Query %			DB	ID	Description
		Match	Length				
C 1	231.6	3.3	659	8	AQ989968	AQ989968 Rfc00652	
C 2	131.4	1.9	621	8	AQ980546	AQ980546 Rfc01339	
C 3	77.6	1.1	315	6	CB041878	CB041878 4008842 B	
C 4	77.6	1.1	315	6	CB042040	CB042040 4007005 B	
C 5	77.6	1.1	349	6	CB041738	CB041738 4006549 B	
C 6	77.6	1.1	359	6	CB041616	CB041616 4006367 B	
C 7	77.4	1.1	314	6	CB061284	CB061284 4011834 B	
C 8	75.8	1.1	300	6	CB061218	CB061218 4011666 B	
C 9	75.8	1.1	317	6	CB059804	CB059804 4003880 B	
C 10	74.4	1.1	354	6	CB041601	CB041601 4006361 B	
C 11	74.2	1.1	302	6	CB062064	CB062064 4012128 B	
C 12	73.6	1.1	475	8	BH439236	BH439236 BOHAC18TF	
C 13	73.6	1.1	754	8	BZ434062	BZ434062 BONFQ34TF	
C 14	72	1.0	425	2	BE323612	BE323612 NF006A09P	
C 15	72	1.0	709	8	BH502634	BH502634 BOHJK46TR	
C 16	72	1.0	736	8	CAY16997	CAY16997 EST641144	
C 17	71.8	1.0	775	6	BZ550668	pacsl-60	
C 18	71.6	1.0	809	7	CF517921	CAP0005_I	
C 19	71.6	1.0	833	7	CF513004	Cabud0005	
C 20	71.2	1.0	632	7	CNI44185	CNI44185 WOUND1_20	
C 21	71.2	1.0	725	7	CNI44663	WOUND1_23	
C 22	70.8	1.0	1218	8	BZ575772	mshs_462.	
C 23	70.4	1.0	682	5	BQ956404	QGB3g04.Y	
C 24	70	1.0	453	1	AI442185	AS58c11.Y	

DEFINITION 4006842 BARC-EMBRAPA 315BOV Bos indicus cDNA clone 315BOV_1E16
 Unknown, mRNA sequence.
 ACCESSION CB041878
 VERSION CB041878.1 GI:27761123
 KEYWORDS EST.
 SOURCE Bos indicus (zebu)
 ORGANISM Bos indicus (zebu)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Ruminantia; Cetartiodactyla; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 315)
 AUTHORS da Mota, A.F., Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Wood, D.L., Capuco, A.V., Brito, M.A.P., Martinez, M.L., Connor, E.E., Machado, M.A. and Coutinho, L.L.
 TITLE Construction and Characterization of cDNA Libraries Generated from Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos indicus) Cattle
 JOURNAL Unpublished (2002)
 COMMENT Contact: Adilson F. da Mota
 Gene Evaluation and Mapping Laboratory
 USDA, ARS, Animal and Natural Resources Institute
 Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048456
 Fax: 3015048414
 Email: amota@cnpgl.embrapa.br
 Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt - -trim fasta. Vector identified by cross match using options -minmatch 12 -minscore 18
 PCR Primers
 FORWARD: GTTTCCTCCAGTCACGCGTTG
 BACKWARD: TGAGCGGATACAAATTCACACG
 Plate: 1 row: E column: 16
 Seq primer: GTTTCCTCCAGTCACGCGTTG
 High quality sequence stop: 315.
 FEATURES
 source
 1. .315
 /organism="Bos indicus"
 /mol_type="mRNA"
 /strain="Brazilian Dairy Gir"
 /db_xref="taxon:9915"
 /clone="315BOV_1E16"
 /sex="female"
 /tissue_type="mammary"
 /cell_type="epithelium"
 /dev_stage="involuting"
 /lab_host="K-12"
 /clone_lib="BARC-EMBRAPA 315BOV"
 /note="Organ: mammary; Vector: pUC118; Site 1: HincII; Site 2: HincII; This mammary-derived cDNA library was created as part of a collaborative project between the ARS Gene Evaluation and Mapping Laboratory and the EMBRAPA Dairy Cattle Research Center under the sponsorship of USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA extracted on 6/8/02, RT with Superscript II at 37 deg C annealing temperature, PCR with 16mer AM5."
 ORIGIN
 Query Match 1.1%; Score 77.6; DB 6; Length 315;
 Best Local Similarity 56.2%; Pred. No. 2.5e-09;
 Matches 146; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
 QY 6313 CGTCATTCCGGACCAATCCATCTGCGAGCGGCCACATGCTTGCACACCGACATAATACG 6372
 DB 306 CGCACTTCCATCGCATCGGACATCCAAAGCCACAGCGTGAACCCGAGAGCCACAAA 247
 QY 6373 CACAATCGCCAGTACCGCAACACCCAGCATAAAGATCGTTGAGCTCAACTTACCGCT 6432
 DB 246 GATGATGCGCGGGCGGGAACACCCGGAACAGCGTCTTCTCCAGCAGCTCGTC 187
 QY 6433 GTCCGGTTTCATGTCGCGACAGATGCCAGCGCCATCCCAACCGTCGATGATTTATG 6492
 DB 186 GTCCGGCTCATGATGCTGCGGTGCCAGGTCCACAGCAGGCGCGTGCATGACATAACGATG 127

QY 6493 CGACAGCGCGCTACGATTTCATCACCACCGGTTGCTCCACACAGATAGCAGCTCCA 6552
 DB 126 CATCGCCCAAGCGAAGACCTCCATCCCGAGGAGTGCGGAGGACAGCAGGATCATCGT 67
 QY 6553 TAACACAGACGATTTGTTCTGTC 6572
 DB 66 CGGCCAGGACATGACCCATC 47
 RESULT 4
 CB042040/c
 LOCUS CB042040.1
 DEFINITION 4007005 BARC-EMBRAPA 315BOV Bos indicus cDNA clone 315BOV_1P23
 Unknown, mRNA sequence.
 ACCESSION CB042040
 VERSION CB042040.1 GI:27761285
 KEYWORDS EST.
 SOURCE Bos indicus (zebu)
 ORGANISM Bos indicus (zebu)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Ruminantia; Cetartiodactyla; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 315)
 AUTHORS da Mota, A.F., Sonstegard, T.S., Van Tassell, C.P., Matukumalli, L.K., Wood, D.L., Capuco, A.V., Brito, M.A.P., Martinez, M.L., Connor, E.E., Machado, M.A. and Coutinho, L.L.
 TITLE Construction and Characterization of cDNA Libraries Generated from Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos indicus) Cattle
 JOURNAL Unpublished (2002)
 COMMENT Contact: Adilson F. da Mota
 Gene Evaluation and Mapping Laboratory
 USDA, ARS, Animal and Natural Resources Institute
 Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048456
 Fax: 3015048414
 Email: amota@cnpgl.embrapa.br
 Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt - -trim fasta. Vector identified by cross match using options -minmatch 12 -minscore 18
 PCR Primers
 FORWARD: GTTTCCTCCAGTCACGCGTTG
 BACKWARD: TGAGCGGATACAAATTCACACG
 Plate: 1 row: P column: 23
 Seq primer: GTTTCCTCCAGTCACGCGTTG
 High quality sequence stop: 315.
 FEATURES
 Location/Qualifiers
 1. .315
 /organism="Bos indicus"
 /mol_type="mRNA"
 /strain="Brazilian Dairy Gir"
 /db_xref="taxon:9915"
 /clone="315BOV_1P23"
 /sex="female"
 /tissue_type="mammary"
 /cell_type="epithelium"
 /dev_stage="involuting"
 /lab_host="K-12"
 /clone_lib="BARC-EMBRAPA 315BOV"
 /note="Organ: mammary; Vector: pUC118; Site 1: HincII; Site 2: HincII; This mammary-derived cDNA library was created as part of a collaborative project between the ARS Gene Evaluation and Mapping Laboratory and the EMBRAPA Dairy Cattle Research Center under the sponsorship of USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA extracted on 6/8/02, RT with Superscript II at 37 deg C annealing temperature, PCR with 16mer AM5."
 ORIGIN
 Query Match 1.1%; Score 77.6; DB 6; Length 315;
 Best Local Similarity 56.2%; Pred. No. 2.5e-09;
 Matches 146; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

6313 CGTATTCGGGACCAATCCACTGCGAGGGCCACATGCTTCCACACCGACATAATCAG 6372
|||
306 CGCACTTCCGATCGGCGATCCGAAGGCCAGACGTCGAACCCGAGAGCCACAAA 247
|||
6373 CACAATCGCCAGTACCGGAAACACACCGCATATAAGATCGTTGAGCTCAAACTTACCGCT 6432
|||
246 GATGATGCGCGGGCGGCGGAAACACCGCGGACAGGTGCTTCTTCCAGCAGCTGTC 187
|||
6433 GTGCGGTTTCATGTCGACAGATGCGAGCCCGCCATCCCAACCGTGCATGATGATTATG 6492
|||
186 GTGCGGCTCATGATGCTGCGGTGCCAGGTCCACAGACGCGGTGCATGACATAACGATG 127
|||
6493 CGACAGCGCGGTACGATTTCCATCACACCGACCGTTGCGGACCAAGATAGACGTTCCA 6552
|||
126 CATCGCCCAAGCGAAGACCTCCATCCCGAGGAGTGGCGGAGAACAGCAGGATCATCGT 67
|||
6553 TAACCAGAGCATGTTTCGTC 6572
|||
66 CGGCCAGGACATGACCCATC 47
|||

RESULT 5
CB041738/c
LOCUS
DEFINITION
4006549 BARC-EMBRAPA 314BOV Bos indicus cdna clone 314BOV_1D23
Unknown, mRNA sequence.

ACCESSION
CB041738
VERSION
CB041738.1 GI:27760983
KEYWORDS
EST.
SOURCE
Bos indicus (zebu)
ORGANISM
Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

REFERENCE
1 (bases 1 to 349)
da Mota,A.F., Sonstegard,T.S., Van Tassel,C.P., Matukumalli,L.K.,
Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,
Machado,M.A. and Coutinho,L.L.
Construction and Characterization of cDNA Libraries Generated from
Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos
indicus) Cattle

JOURNAL
COMMENT
Unpublished (2002)
Contact: Adilson F. da Mota
Gene Evaluation and Mapping Laboratory
USDA, ARS, Animal and Natural Resources Institute
Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA
Tel: 3015048456
Fax: 3015048414

Email: amota@cnpgl.embrapa.br
Single pass sequencing. Bases called and trimmed with phred.
0.000925 using options -trim alt -trim fasta. Vector identified
by cross match using options -minmatch 12 -minscore 12

PCR Primers
FORWARD: GTTTCCTCCAGTCAGCGTTG
BACKWARD: TGAGCGGATACAAATTTCCACAG
Plate: 1 row: D column: 23
Seq primer: GTTTCCTCCAGTCAGCGTTG
High quality sequence stop: 349.

FEATURES
source
1..349
/organism="Bos indicus"
/mol_type="mRNA"
/strain="Brazilian Dairy Gir"
/db_xref="taxon:9915"
/clone="314BOV_1D23"
/sex="female"
/tissue_type="test cistern"
/cell_type="epithelium"
/dev_stage="involutus"
/lab_host="K-12"
/clone_lib="BARC-EMBRAPA 314BOV"
/note="Organ: mammary; Vector: pUC118; Site 1: HincII;
Site 2: HincII; This mammary-derived cdna library was

created as part of a collaborative project between the ARS
Gene Evaluation and Mapping Laboratory and the EMBRAPA
Dairy Cattle Research Center under the sponsorship of
USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA
extracted on 6/8/02, RT with Superscript II at 37 deg C
annealing temperature, PCR with 16mer AM5 only."

ORIGIN

Query Match 1.1%; Score 77.6; DB 6; Length 349;
Best Local Similarity 56.2%; Pred. No. 2.6e-09;
Matches 146; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 6313 CGTCAATTCGGGACCAATCCACTGCGAGGGCCACATGCTTCCACACCGACATAAATCAG 6372
DB 340 CGCACTTCCGATCGGCGATCCGAAGGCCAGACGTCGAACCCGAGAGCCACAAA 281
QY 6373 CACAATCGCCAGTACCGGAAACACACCGCATATAAGATCGTTGAGCTCAAACTTACCGCT 6432
DB 280 GATGATGCGCGGGCGGCGGAAACACCGCGGACAGGTGCTTCTTCCAGCAGCTGTC 221
QY 6433 GTGCGGTTTCATGTCGACAGATGCGAGCCCGCCATCCCAACCGTGCATGATTATG 6492
DB 220 GTGCGGCTCATGATGCTGCGGTGCCAGGTCCACAGACGCGGTGCATGACATAACGATG 161
QY 6493 CGACAGCGCGGTACGATTTCCATCACACCGACCGTTGCGGACCAAGATAGACGTTCCA 6552
DB 160 CATCGCCCAAGCGAAGACCTCCATCCCGAGGAGTGGCGGAGAACAGCAGGATCATCGT 101
QY 6553 TAACCAGAGCATGTTTCGTC 6572
DB 100 CGGCCAGGACATGACCCATC 81

RESULT 6

CB041616
LOCUS
DEFINITION
4006367 BARC-EMBRAPA 314BOV Bos indicus cdna clone 314BOV_1F09
Unknown, mRNA sequence.

ACCESSION
CB041616
VERSION
CB041616.1 GI:27760861
KEYWORDS
EST.
SOURCE
Bos indicus (zebu)
ORGANISM
Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.

REFERENCE
1 (bases 1 to 359)
da Mota,A.F., Sonstegard,T.S., Van Tassel,C.P., Matukumalli,L.K.,
Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,
Machado,M.A. and Coutinho,L.L.
Construction and Characterization of cDNA Libraries Generated from
Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos
indicus) Cattle

JOURNAL
COMMENT
Unpublished (2002)
Contact: Adilson F. da Mota
Gene Evaluation and Mapping Laboratory
USDA, ARS, Animal and Natural Resources Institute
Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA
Tel: 3015048456
Fax: 3015048414

Email: amota@cnpgl.embrapa.br
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt -trim fasta. Vector identified
by cross match using options -minmatch 12 -minscore 12

PCR Primers
FORWARD: GTTTCCTCCAGTCAGCGTTG
BACKWARD: TGAGCGGATACAAATTTCCACAG
Plate: 1 row: F column: 09
Seq primer: GTTTCCTCCAGTCAGCGTTG
High quality sequence stop: 359.

FEATURES

source
1..359
/organism="Bos indicus"

Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt - -trim fasta. Vector identified by cross_match using options -mismatch 12 -minscore 12

PCR Primers

FORWARD: GTTTTCCAGTCACGACGTTG

BACKWARD: TGAGGGATACAAATTTACACAG

Plate: 1 row: H column: 20

Seq primer: GTTTTCCAGTCACGACGTTG

High quality sequence stop: 314.

Location/Qualifiers

1. 314

/organism="Bos indicus"

/mol_type="mRNA"

/strain="Brazilian Dairy Gir"

/db_xref="taxon:9915"

/clone="338BOV_1H20"

/sex="female"

/tissue_type="teat cistern and alveoli"

/cell_type="multiple"

/dev_stage="involved"

/lab_host="K-12"

/clone_lib="BARC-EMBRAPA 338BOV"

/note="Organ: mammary; Vector: pUC 118; Site 1: HincII;

Site 2: HincII; This mammary-derived cDNA library was

created as part of a collaborative project between the ARS

Gene Evaluation and Mapping Laboratory and the EMBRAPA

Dairy Cattle Research Center under the sponsorship of

USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA

extracted on 6/8/02, RT with Superscript II at 37 deg C

annealing temperature. Cow AM1."

FEATURES

source

ORIGIN

Query Match 1.1%; Score 77.4; DB 6; Length 314;
Best Local Similarity 57.1%; Pred. No. 2.8e-09;
Matches 141; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

6326 CCAATTCCTGTCAGCGGCCACATGCTTGCACACCGACATATACGACATCGCCAGT 6385
|||||
4 CCGATCGGCTGTCATCCAGGCCAGAGCGTGCACACCGGAGCCACAAAGATGATGCGCGGG 63

6386 ACAGCAACACACCGCATAAAGATCGTTGAGCTCAAACTTACCGGTGTCGGTTCATGG 6445
|||||
64 GGGCGGAAACACCGGACAGGTCGTTCTTCCAGCACGTCGTCGCGGCTCATGA 123

6446 TCGACAGATGCCAGCCCATCCCAACCGTCGATGATGATTTATGCGACACGCGCGT 6505
|||||
124 TGGTCGCGTGCAGGTCACACGAGCGCGTGCATGACATACGATGCATCGCCATCG 183

6506 ACATTTCCATCACCACCGAGGTTGCCACACAGATAGCAGCTTCCATACGACGATTT 6565
|||||
184 AAGACCTCCATCCCGAGGAAGTGGCGAGGACAGCAGGATCATCTCGCGCCAGGACATG 243

6566 GTTCGTC 6572
244 ACCATC 250

RESULT 8

CB061218/c

LOCUS

4011666 BARC-EMBRAPA 338BOV Bos indicus cDNA clone 338BOV_1L08

DEFINITION

Unknown, mRNA sequence.

ACCESSION

CB061218

VERSION

CB061218.1 GI:27799505

KEYWORDS

EST.

SOURCE

Bos indicus (zebu)

ORGANISM

Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovinae; Bos.

REFERENCE

1 (bases 1 to 300)

da Mota,A.F., Sonstegard,T.S., Van Tassell,C.P., Matukumalli,L.K.,

Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,

/mol_type="mRNA"

/strain="Brazilian Dairy Gir"

/db_xref="taxon:9915"

/clone="314BOV_1F09"

/sex="female"

/tissue_type="teat cistern"

/cell_type="epithelium"

/dev_stage="involved"

/lab_host="K-12"

/clone_lib="BARC-EMBRAPA 314BOV"

/note="Organ: mammary; Vector: pUC118; Site 1: HincII;

Site 2: HincII; This mammary-derived cDNA library was

created as part of a collaborative project between the ARS

Gene Evaluation and Mapping Laboratory and the EMBRAPA

Dairy Cattle Research Center under the sponsorship of

USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA

extracted on 6/8/02, RT with Superscript II at 37 deg C

annealing temperature. PCR with 16mer AMS only."

ORIGIN

Query Match 1.1%; Score 77.6; DB 6; Length 359;
Best Local Similarity 56.2%; Pred. No. 2.6e-09;
Matches 146; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

6313 CGTCATTCGGCACCACATCCATGCGAGCGCCACATGCTTGCACACCGACATATACG 6372
|||||
20 CGCACTTCCATCCGATCGGCAGCATCCAAAGCCAGAGCGTGCACCCGAGGCCACAAA 79

6373 CACAATCGCCAGTACCGCAACACACCGCATAAAGATGTTGAGCTCAAACTTACCGT 6432
|||||
80 GATGATGCGCGGGCGGGGAAACCGGACAGGTCGTTCTTCCAGCAGCTGCTC 139

6433 GTCCGGTTTCATGTCGACAGATGCCAGCCCATCCCAACCGTGCATGATGATTTATG 6492
|||||
140 GTCCGGTTCATGATGCTGCGGTGCCAGTCCACAGCAGCGCGTGCATGACATACGATG 199

6493 CGACAGCGCGTACGATTTCCATCACCACCGTTCGCCAACAGATAGCAGTTCCA 6552
|||||
200 CATCGCCCAAGCAAGACCTTCCATCCCGAGGAGTGGCGAGGACAGCAGGATCATCGT 259

6553 TAACACGAGCATTTGTCGTC 6572
|||||
260 CGGCCAGGATACCCATC 279

RESULT 7

CB061284

LOCUS

4011834 BARC-EMBRAPA 338BOV Bos indicus cDNA clone 338BOV_1H20

DEFINITION

Unknown, mRNA sequence.

ACCESSION

CB061284

VERSION

CB061284.1 GI:27799571

KEYWORDS

EST.

SOURCE

Bos indicus (zebu)

ORGANISM

Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovinae; Bos.

REFERENCE

1 (bases 1 to 314)

da Mota,A.F., Sonstegard,T.S., Van Tassell,C.P., Matukumalli,L.K.,

Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,

Machado,M.A. and Coutinho,L.B.

Construction and Characterization of cDNA Libraries Generated from

Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos

indicus) Cattle

Unpublished (2002)

Contact: Adilson F. da Mota

Gene Evaluation and Mapping Laboratory

USDA, ARS, Animal and Natural Resources Institute

Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA

Tel: 3015048456

Fax: 3015048414

Email: amota@cnpgl.embrapa.br

1 (bases 1 to 314)

da Mota,A.F., Sonstegard,T.S., Van Tassell,C.P., Matukumalli,L.K.,

Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,

Machado,M.A. and Coutinho,L.B.

Construction and Characterization of cDNA Libraries Generated from

Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos

indicus) Cattle

Unpublished (2002)

Contact: Adilson F. da Mota

Gene Evaluation and Mapping Laboratory

USDA, ARS, Animal and Natural Resources Institute

Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA

Tel: 3015048456

Fax: 3015048414

Email: amota@cnpgl.embrapa.br

TITLE	JOURNAL	COMMENT
Machado, M.A. and Coutinho, L.L. Construction and Characterization of cDNA Libraries Generated from Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos indicus) Cattle Unpublished (2002)		Contact: Adilson F. da Mota Gene Evaluation and Mapping Laboratory USDA, ARS, Animal and Natural Resources Institute Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA Tel: 3015048456 Fax: 3015048414 Email: amota@cnpgl.embrapa.br Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt ' ', -trim fasta. Vector identified by crossmatch using options -minmatch 12 -minscore 12 PCR primers FORWARD: GTTTTCCCAAGTCACGCGTTG BACKWARD: TGAGCGGATAACAAATTTCACACAG Plate: 1 row: 1 column: 08 Seq primer: GTTTTCCCAAGTCACGCGTTG High quality sequence stop: 300.

FEATURES

```

ORIGIN
Query Match 1.1%; Score 75.8; DB 6; Length 300;
Best Local Similarity 56.7%; Pred. No. 7.7e-09;
Matches 140; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
annealing temperature: 50.0;

```

RESULT	9
CB059804/c	
LOCUS	CB059804
DEFINITION	4009880 BARC-EMBRAPA 323BOV Bos indicus cDNA clone 323BOV 1B22
	317 bp mRNA linear EST 17-JAN-2003

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	REFERENCE AUTHORS	TITLE	JOURNAL COMMENT
--	----------------------	-------	--------------------

FEATURES source

ORIGIN

	Query Match	1.1%;	Score 75.8;	DB 6;	Length 317;
	Best Local Similarity	56.7%;	Pred. No. 7.8e-09;		
	Matches 140;	Conservative	0;	Mismatches 107;	Indels 0;
	Gaps	0;			
Qy	6326	CCAAATCCACATGCGAGCGGCACATGCTTGCACACCGACATAAATCAGACAAATCGCCAGT			6385
Db	311	CCGATCGCGAGCATCCAAAGGCCAGAGTGTCAACCCGAGAGCCACAAGATGATGGCGCGG			252
Qy	6386	ACCGCAAAACACCAACCGCATAAAGATCGTTGAGCTCAAACCTTACCGCTGTGCGGTTTCATGG			6445
Db	251	CGCGGAAACACCCGCGACAGTCTGTTCTTCTCCAGCAGCGTCGTCGTGCGGCTCATGA			192
Qy	6446	TGGACAGATGCCAGCGCCCATCCCCAAACCGTGCATGATGTATTTATGCCAGACGCGCGCT			6505
Db	191	TGGCTGGGTTGCCAGGTGCCACAGACGGCGGTGATGACATAACGATGCATCGGCCAAGCG			132

QY 6506 ACGATTTCATCACCACCGGTTGCCACAGTAAGACAGTTCATTAACACGAGCATT 6565
 Db 131 AAGACTTCATCCCAAGGAGGTGGCGAGAAACAGCAGGATCGTCGTGGCCAGGACATG 72
 QY 6566 GTTTCGTC 6572
 Db 71 ACCATC 65

RESULT 10
 CB041601/c 354 bp mRNA linear EST 15-JAN-2003
 LOCUS BARC-EMBRAPA 314BOV Bos indicus cDNA clone 314BOV_1E09
 DEFINITION Unknown, mRNA sequence.

ACCESSION CB041601
 VERSION CB041601.1 GI:27760846
 KEYWORDS EST.
 SOURCE Bos indicus (zebu)
 ORGANISM Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.

REFERENCE 1 (bases 1 to 354)
 AUTHORS da Mota,A.F., Sonstegard,T.S., Van Tassel,C.P., Matukumalli,L.K.,
 Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,
 Machado,M.A. and Coutinho,L.L.
 TITLE Construction and Characterization of cDNA Libraries Generated from
 Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos
 indicus) Cattle

JOURNAL Unpublished (2002)
 COMMENT Contact: Adilson F. da Mota
 Gene Evaluation and Mapping Laboratory
 USDA, ARS, Animal and Natural Resources Institute
 Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048456
 Fax: 3015048414
 Email: amota@npgl.embrapa.br
 Single pass sequencing. Bases called and trimmed with phred
 0.000925 using options -trim_alt '' -trim_fasta. Vector identified
 by cross match using options -minmatch 12 -minscore 12
 PCR Primers
 FORWARD: GTTTTCCAGTCACGACGTTG
 BACKWARD: TGAGCGGATACAAATTTCACAG
 Plate: 1 row: E column: 09
 Seq primer: GTTTTCCAGTCACGACGTTG
 High quality sequence stop: 354.

FEATURES
 source
 1..354
 /organism="Bos indicus"
 /mol_type="mRNA"
 /strain="Brazilian Dairy Gir"
 /db_xref="taxon:9915"
 /clone="314BOV_1E09"
 /sex="female"
 /tissue_type="teat cistern"
 /cell_type="epithelium"
 /dev_stage="involved"
 /lab_host="K-12"
 /clone_lib="BARC-EMBRAPA 314BOV"
 /note="Organ: mammary; Vector: pUC118; Site 1: HincII;
 Site2: HincII; This mammary-derived cDNA library was
 created as part of a collaborative project between the ARS
 Gene Evaluation and Mapping Laboratory and the EMBRAPA
 Dairy Cattle Research Center under the sponsorship of
 USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA
 extracted on 6/8/02, RT with Superscript II at 37 deg C
 annealing temperature, PCR with 16mer AWS only."

ORIGIN

Query Match 1.1%; Score 74.4; DB 6; Length 354;
 Best Local Similarity 55.4%; Pred. No. 1.9e-08;
 Matches 144; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 6313 CGTCAATTCGGCACCAATCCATCGAGCGGCCACATGCTTGCACACCGACATAAATCAG 6372
 Db 345 CGCACTTCCATCGCCGATCGGAGCATCCAAGGCCAGAGCTGCAACCCGAGAGCCACAA 286
 QY 6373 CACAATCGCAGTACGCAAAACACACCGCATATAAGATCGTTGAGCTCAAACCTTACCGCT 6432
 Db 285 GATGATGGCGGGGGCGGAAACACCGCGAACAGGTGCTCTTCTCCAGCAGCTCGTC 226
 QY 6433 GTGCGGTTTCATGTCGACAGATGCCAGCCCATCCCCAACCGTGCATGATGATTTATG 6492
 Db 225 GTGCGGCTCATGATGCTGCGGTGCCAGGTCCACAGCAGCGGTGCATGACATAACGATG 166
 QY 6493 GCACAGCGCGCTACGATTTCCATCACACACCGGTTGCCAACAGATAGACAGCTTCCA 6552
 Db 165 CATGCCCCAAAGCGAAGACCTCGTCCCCAGGAAGTGGCGGAAACAGCAGGATCATCGT 106
 QY 6553 TAAACAGAGCATTTGTTGCTC 6572
 Db 105 CGGCCAGGACATGACCCATC 86

RESULT 11
 CB062064/c 302 bp mRNA linear EST 17-JAN-2003
 LOCUS BARC-EMBRAPA 339BOV Bos indicus cDNA clone 339BOV_1114
 DEFINITION Unknown, mRNA sequence.

ACCESSION CB062064
 VERSION CB062064.1 GI:27800351
 KEYWORDS EST.
 SOURCE Bos indicus (zebu)
 ORGANISM Bos indicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.

REFERENCE 1 (bases 1 to 302)
 AUTHORS da Mota,A.F., Sonstegard,T.S., Van Tassel,C.P., Matukumalli,L.K.,
 Wood,D.L., Capuco,A.V., Brito,M.A.P., Martinez,M.L., Connor,E.E.,
 Machado,M.A. and Coutinho,L.L.
 TITLE Construction and Characterization of cDNA Libraries Generated from
 Mammary Gland Tissues of Holstein (Bos taurus) and Gir (Bos
 indicus) Cattle

JOURNAL Unpublished (2002)
 COMMENT Contact: Adilson F. da Mota
 Gene Evaluation and Mapping Laboratory
 USDA, ARS, Animal and Natural Resources Institute
 Bldg. 200 Rm3 BARC-East, Beltsville, MD 20705, USA
 Tel: 3015048456
 Fax: 3015048414
 Email: amota@npgl.embrapa.br
 Single pass sequencing. Bases called and trimmed with phred
 0.000925 using options -trim_alt '' -trim_fasta. Vector identified
 by cross match using options -minmatch 12 -minscore 12
 PCR Primers
 FORWARD: GTTTTCCAGTCACGACGTTG
 BACKWARD: TGAGCGGATACAAATTTCACAG
 Plate: 1 row: 1 column: 14
 Seq primer: GTTTTCCAGTCACGACGTTG
 High quality sequence stop: 302.

FEATURES
 source
 1..302
 /organism="Bos indicus"
 /mol_type="mRNA"
 /strain="Brazilian Dairy Gir"
 /db_xref="taxon:9915"
 /clone="339BOV_1114"
 /sex="female"
 /tissue_type="teat cistern and alveoli"
 /cell_type="multiple"
 /dev_stage="involved"
 /lab_host="K-12"
 /clone_lib="BARC-EMBRAPA 339BOV"
 /note="Organ: mammary; Vector: pUC 19; Site 1: SmaI;
 Site_2: SmaI; This mammary-derived cDNA library was

created as part of a collaborative project between the ARS Gene Evaluation and Mapping Laboratory and the EMBRAPA Dairy Cattle Research Center under the sponsorship of USDA, ARS/EMBRAPA-LABEX program in animal genomics. RNA extracted on 6/8/02, RT with Superscript II at 37 deg C annealing temperature. Cow AW1."

ORIGIN

Query Match 1.1%; Score 74.2; DB 6; Length 302;
Best Local Similarity 56.3%; Pred. No. 2.1e-08;
Matches 139; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
QY 6326 CCAATCCATGCGAGCGGCACATGCTTCACACCGACATAAATACGACAAATCGCCATT 6385
DB 296 CCGATCGGAGCATCAAGCGCAGATGTGCAACCCGAGAGCCACAAAGATGATGCGCGG 237
QY 6386 ACCGCAAAACACACCGCATTAAGATCGTTGAGCTCAAACTTACCGCTGTCGCGTTTCATGG 6445
DB 236 GCGGCGAAACACACCGCATTAAGATCGTTGAGCTCAAACTTACCGCTGTCGCGTTTCATGG 177
QY 6446 TCGCAGCATGCGAGCGCCCATCCCAACCGTGCATGATGTATTATGCGACAGCGCGCT 6505
DB 176 TGGCTCGGTGCAGGTGCACAGCAGCGCGTGCATGACATACGATCGATCGCCCAAGCG 117
QY 6506 AGATTTCATCACCACACCGTGTGCAACAGATGAAGACATTCATTAACACGAGCAT 6565
DB 116 AAGACCTCCATCCCGAGGAGGTGGCGAGAACAGCAGGATCATCGTGGCCGAGGACATG 57
QY 6566 GTTCGTC 6572
DB 56 ACCATC 50

RESULT 12

BH439236/c
LOCUS BH439236 475 bp DNA linear GSS 12-DEC-2001
DEFINITION BOHAC18TF BOHA Brassica oleracea genomic clone BOHAC18, genomic survey sequence.
ACCESSION BH439236
VERSION BH439236.1 GI:17624950
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 475)
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BOHAC18TR
Contact: Chris Town

FEATURES

source
1..475
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BOHA"
/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 1.1%; Score 73.6; DB 8; Length 475;
Best Local Similarity 53.9%; Pred. No. 3.9e-08;
Matches 151; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 559 GCCTGTGAGTGAATGTCACCGCGCATGCTGTGATTTGGATGACATTCCTCGATG 618
DB 194 GCTTGGCGGTGAATGATCCACAGCATGTCGCTGATCCACGACGATCTCCCGTGTATG 253
QY 619 GATACGCGCAGATCGCTGCTGCCCTACCGTGCATCGCGAATTTGGTGAACACG 678

Best Local Similarity 53.9%; Pred. No. 3.5e-08;
Matches 151; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 559 GCCTGTGAGTGAATGTCACCGCGCATGCTGTGATTTGGATGACATTCCTCGATG 618
DB 378 GCTTGGCGGTGAATGATCCACAGCATGTCGCTGATCCACGACGATCTCCCGTGTATG 319
QY 619 GATACGCGCAGATCGCTGCTGCCCTACCGTGCATCGCGAATTTGGTGAACACG 678
DB 318 GACACGACGACCTCCGCGCGGAAACCCACCAACCAAGTGTTCGGCGAAGACGTC 259
QY 679 GCGATTTCGCGCCCATCGCGCTGTAGCCCGCGCATTTGAAGTGAATTCCTCGATG 738
DB 258 GCGGTTTTCGCGGAGACGCGCTTTTATCGTTTCGCTTTCGCTGAGCACTTCGCGTCGACGTCG 199
QY 739 GCTTTCCTGCGCATATAAATCTGAAGCATGCTGAACCTCTCCGCTGCCGTCGCGCTG 798
DB 198 GCGGTGCTTCGCGCGAGGGTGTAGACGATCGGGAGCTGGCGAGAGCCGTTGGATCA 139
QY 799 CAGGCGCTTAGTGCAGAGGCAATTCAGGATCTGCACGACG 838
DB 138 GAAGGCTTGTGGCGGTTCAGATTGTGGATATCAGCAGCG 99

RESULT 13

BZ434062
LOCUS BZ434062 754 bp DNA linear GSS 13-DEC-2002
DEFINITION BONFQ34TF BO 1.6.2_KB tot Brassica oleracea genomic clone BONFQ34, genomic survey sequence.
ACCESSION BZ434062
VERSION BZ434062.1 GI:26683951
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 754)
AUTHORS Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other GSSs: BONFQ34TR
Contact: Chris Town

FEATURES

source
1..754
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BONFQ34"
/clone_lib="BO 1.6.2_KB tot"
/note="Vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 1.1%; Score 73.6; DB 8; Length 754;
Best Local Similarity 53.9%; Pred. No. 3.9e-08;
Matches 151; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 559 GCCTGTGAGTGAATGTCACCGCGCATGCTGTGATTTGGATGACATTCCTCGATG 618
DB 194 GCTTGGCGGTGAATGATCCACAGCATGTCGCTGATCCACGACGATCTCCCGTGTATG 253
QY 619 GATACGCGCAGATCGCTGCTGCCCTACCGTGCATCGCGAATTTGGTGAACACG 678

```

Db      254  GACAACGACGACCTCCGCGCGGAAACCCACCAACCAACCAAGTGTTCGCGAAGACGTC 313
QY      679  GCGAATTCCTCCGCCCATCCGCTGTAGCCGCGCATTTGAAGTGAATGCCATTGCACCC 738
Db      314  GCGGTTTAAAGCCGAGACGCGCTTTATCTTCGCGTTTGAGCAGCTTGGCGTCGACGTCG 373
QY      739  GGTTCCTCCGATACATAAATCTGAAGGATGCTGAACCTCTCCGCTCCGCTCGGCGCTG 798
Db      374  GGGGTGGCTCCGCGGAGGAGGTGTAGAGGATCGGGAGCTGGCGAGAGCCGCTTGGATCA 433
QY      799  CAGGGCTTAGTCAAGGCAATTCACAGGATCTGCAGGACG 838
Db      434  GAAGGCTTGTGCGGGTCAGATTGGATATCAGCAGCG 473

RESULT 14
LOCUS   BE323612
DEFINITION BE323612 425 bp mRNA linear EST 21-DEC-2000
ACCESSION NF006A09PL1065 Phosphate starved leaf Medicago truncatula cDNA
VERSION clone NF006A09PL 5', mRNA sequence.
KEYWORDS BE323612
SOURCE BE323612.2 GI:11967172
ORGANISM Medicago truncatula (barrel medic)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 425)
AUTHORS Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
TITLES Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
JOURNAL Expressed Sequence Tags from the Samuel Roberts Noble Foundation
COMMENT Medicago truncatula phosphate-starved leaf library
Unpublished (2000)
On Jul 14, 2000 this sequence version replaced gi:9197389.
Contact: Harrison MJ
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7325
Fax: 580 221 7380
Email: mjharrison@noble.org
Medicago Genome Initiative accession: MGI:S:20177
Insert Length: 818 Std Error: 0.00
Plate: 006 row: A column: 09
Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES
source Location/Qualifiers
1..425
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF006A09PL"
/tissue_type="leaf"
/dev_stage="trifoliolate"
/clone_lib="phosphate starved leaf"
/note="Vector: lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 200M potassium phosphate. RNA was prepared from above ground tissues."
ORIGIN
Query Match 1.0%; Score 72; DB 2; Length 425;
Best Local Similarity 62.0%; Pred. No. 9.2e-08;
Matches 114; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 558 CGCTGTGTCAGTGAATGGTGCACCGGCATCGCTGATTCGTGATGACATTCCTCCGAT 617
Db 233 CGCTGCGCGGTGAATGATCCACAGATGCTCTCTCCACGATGACCTTCCTTGAT 292
QY 618 GGATAACGCGCAGATCGCTGCTGGTCCCTACCGTGCATCGCGAATTTGGTGAACGCT 677

```

```

Db      293  GGATAACGATGATCTCCGCGAGGTAAACCTACAAACCAAGTCTTCGGAGAAGACGT 352
QY      678  GCGCATTTCTCGCGGCATCCGCTGTAGCCGCGCATTTGAAGTGAATGCCATTGCACCC 737
Db      353  TCGTGTCTTCGCGCGGAGATGCTCTCTCTCGCTTTGCTTCGAACATATTGCCGTCTCCAC 412
QY      738  CGGT 741
Db      413  CGTT 416

RESULT 15
LOCUS   BH502634
DEFINITION BH502634 709 bp DNA linear GSS 13-DEC-2001
ACCESSION BOHJK46TR BOHJ Brassica oleracea genomic clone BOHJK46, genomic
VERSION BH502634 survey sequence.
KEYWORDS BH502634.1 GI:17710731
SOURCE GSS.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 709)
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
TITLES Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: BOHJK46TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
source Location/Qualifiers
1..709
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHJK46"
/clone_lib="BOHJ"
/note="Vector: pHS01; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into pHS01 using BstXI linkers"
ORIGIN
Query Match 1.0%; Score 72; DB 8; Length 709;
Best Local Similarity 50.2%; Pred. No. 1.1e-07;
Matches 239; Conservative 0; Mismatches 225; Indels 12; Gaps 2;

QY 559 GCCTGTGTCAGTGAATGGTGCACCGGCATCGCTGATTCGTGATGACATTCCTCGATG 618
Db 87 GCTTGGCGGTTCGATGATTCACACAGATGTCATCATCAAGACGATCTCTCTTCGATG 146
QY 619 GATACGCGCAGATCGCTGCTGCTGCCCTACCGTGCATCGCGAATTTGGTGAACGCTG 678
Db 147 GACAACGACGACCTCCGCGAGGTAAAGCCACGACCAAAAGTCTTCGGAGAAGCGTC 206
QY 679 GCGATTTCGCGCCCATCGCTGTAGCCCGCATTTGAAGTGAATTCGCAATTCGACCC 738
Db 207 GCGATTCTCTCCGCGCGCGCTCTTAGCTCTGGCTTCGAGCATTTGACGAAGTGAC 266
QY 739 GGTTCCTCCGATACATAAATCTGAAGCGATTCGTGAACCTCTCCGCTCCGCTCGGCTG 798
Db 267 GTGTCGCTGACAGA---ATGTTAGAGCGGTTAAAGAACTCGCGAAGTCTATAGGACG 323
QY 799 CAGGGCTTAGTCAAGGCAATTCAGGATCTGACGAC-----GGCAGCGAGACG 849

```


Db 324 AAAGGCTCGTGGCGGACAAGCGATGGATTGTGAGCAGCGAAGGTTTGGATCAAAACGAC 383
QY 850 CGCAGCCCGGAAGCGATCGCCATGACCAACGAACCTGAATAACAGCGTGTCTTTTCGCGCC 909
Db 384 GTCGGTTGAAGGAGCTTGAGTTTATTACAGTTTCATAAACCGGTTTCGCTGCTTGAGGCT 443
QY 910 ACGCTGCMAATGGCGGCGATTGCGGCTGACGCTTACCGCAGGTTGCGGCAAGACTTAGC 969
Db 444 TCGGCGTTATCGGAGCGGTTATTGGAGGTGTTTCGGAGGAAGGTTGAGAAGGTGAGG 503
QY 970 TTCTTCGCCCGAGATTGGGCCAGGCGTTTCACTGCTCGACGACCTCGCCGACGG 1025
Db 504 AGGTTCCGAGGTGATTGGGTTGTTTTCAGGTGTTGATGATATTTTGGATGG 559

Search completed: January 25, 2005, 10:29:04
Job time : 19364 secs